Maria do Pilar Castillo Portela

Licenciatura em Bioquímica

Structural and functional characterization of the periplasmic cytochrome PpcA from *Geobacter metallireducens*

Dissertação para obtenção do Grau de Mestre em Bioquímica

Orientador: Prof. Doutor Carlos A. Salgueiro, Professor Associado com Agregação, Faculdade de Ciências e Tecnologia, Universidade NOVA de Lisboa

Júri:

Presidente: Prof. Doutor José Ricardo Ramos Franco Tavares
Arguente: Prof. Doutora Maria Alice Santos Pereira
Vogal: Prof. Doutor Carlos Alberto Gomes Salgueiro

FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Setembro 2019

Universidade Nova de Lisboa

Maria do Pilar Castillo Portela

Licenciatura em Bioquímica

Structural and functional characterization of the periplasmic cytochrome PpcA from *Geobacter metallireducens*

Dissertação para obtenção do Grau de Mestre em Bioquímica

Orientador: Prof. Doutor Carlos A. Salgueiro, Professor Associado com Agregação, Faculdade de Ciências e Tecnologia, Universidade NOVA de Lisboa

FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Setembro 2019

Structural and functional characterization of the periplasmic cytochrome PpcA from *Geobacter metallireducens*

"Copyright"

Maria do Pilar Castillo Portela
Faculdade de Ciências e Tecnologia
Universidade NOVA de Lisboa

Structural and functional characterization of the periplasmic cytochrome PpcA from *Geobacter metallireducens*

The results in Chapter 4 were published on the Biomolecular NMR Assignments journal[1] and were presented as a flash presentation in the Modern Methods of Structure Elucidation Course 2019[2]. Part of the results obtained in Chapter 5 were published on The Journal of Physical Chemistry B[3].

[1]P.C. Portela, J.M. Dantas, C.A. Salgueiro, Backbone, side chain and heme resonance assignment of the triheme cytochrome PpcA from *Geobacter metallireducens* in the oxidized state. Biomolecular NMR Assignments (2019).

[2]P.C. Portela, J.M. Dantas, C.A. Salgueiro, Looking for the cofactors: NMR characterization of the triheme cytochrome PpcA from *Geobacter metallireducens* in the oxidized state, 2019.

[2] J.M. Dantas, P.C. Portela, A.P. Fernandes, Y.Y. Londer, X. Yang, N.E.C. Duke, M. Schiffer, P.R. Pokkuluri, C.A. Salgueiro, Structural and functional relevance of the conserved residue V13 in the triheme cytochrome PpcA from *Geobacter sulfurreducens*, The Journal of Physical Chemistry B, 123 (2019) 3050-3060.

The work developed in this thesis was supported by Fundação para a Ciência e Tecnologia (FCT) through the grants: PTDC/BBB-BQB/3554/2014 and PTDC/BIA-BQM/31981/2017. This work was also supported by the Applied Molecular Biosciences Unit – UCIBIO which is financed by national funds from FCT/MCTES (UID/Multi/04378/2019). The NMR spectrometers at FCT NOVA are part of Rede Nacional de RMN (PTNMR), supported by FCT-MCTES (ROTEIRO/0031/2013 - PINFRA/22161/2016) co-funded by FEDER through COMPETE 2020, POCI, and PORL and FCT through PIDDAC.

Agradecimentos

A presença de muitas pessoas, tanto dentro como fora do laboratório, foi essencial para a concretização deste trabalho e aqui deixo o meu agradecimento a todas elas.

Gostaria de agradecer ao Professor Doutor Carlos Salgueiro pela maneira irrepreensível com que desempenhou o seu papel de orientador. Agradeço toda a dedicação, paciência, trabalho de equipa e ensinamentos transmitidos que permitiram que eu crescesse como profissional. A par disto, as sempre sensatas palavras, especialmente em alturas mais difíceis, e a sua disponibilidade para me ouvir fizeram sempre com que a célebre frase de Confúcio se aplicasse "Escolhe um trabalho de que gostes, e não terás que trabalhar nem um dia na tua vida".

À Doutora Marta Silva pela ajuda crucial nas titulações UV-Visível na câmara e nos crescimentos dos mutantes, pelo entusiasmo pelo nosso trabalho bem como pelas grandes conversas (especialmente em dias de titulação!).

À Doutora Leonor Morgado por estar sempre disponível para as minhas dúvidas, pelos sempre proveitosos conselhos bem como pelas conversas diárias.

Aos colegas de laboratório, Mestres Liliana Teixeira, Marisa Ferreira e Tomás Fernandes por tornarem o Lab. 611 um local de trabalho tão divertido quanto de aprendizagem e por sempre estarem disponíveis para me ajudar. Gostaria de agradecer de forma especial à Liliana pelos conhecimentos de PCR transmitidos, essenciais para a concretização dos mutantes desta tese.

À Doutora Joana Dantas pelo apoio em todos os assignments e cálculos que levaram à estrutura do PpcA, bem como por toda a ajuda em muitas outras dúvidas que foram surgindo.

Aos meus Pais, por fazerem com que seja quem sou. Por todo o carinho, por toda a compreensão nas muitas horas de trabalho fora de casa e pelo ânimo durante a escrita desta tese. Espero que com este documento possam ver como gosto daquilo que faço! Aos meus irmãos, por todos os risos e brincadeiras. Pelas certeiras piadas do Luís e pelos calorosos abraços da Victória. A todos os meus avós e restante família por me apoiarem diariamente. Um agradecimento especial aos meus Abuelos *por la pequeña gran casa Agarimo con su hermoso jardín*, fundamental na escrita desta tese.

Ao Paulo, por estares sempre perto desde sempre! Pela tua paciência, carinho, compreensão e apoio incondicional todos os dias tanto nos bons como nos maus momentos. Por tudo o resto que é impossível de materializar em palavras mas que é "essencial mas invisível aos olhos". Por tornares teus também os meus sonhos!

Aos meus amigos, por me acompanharem nesta jornada desde há seis anos. Obrigada por tudo!

Abstract

The bacterium *Geobacter metallireducens* (*Gm*) is capable of transferring electrons to the cell's exterior and thus reduce extracellular electron acceptors. Because of this, *Gm* has been used for electricity harvesting upon its association with electrodes and for bioremediation of contaminated waters with Cr(VI) and U(VI), for example, which serve as extracellular acceptors. The triheme *c*-type cytochrome PpcA from *Gm* is abundant in the periplasm and crucial for bridging the electron transfer between the cytoplasm and the cell's exterior. It shares 80% of identity with the well-characterized PpcA from *Geobacter sulfurreducens* (*Gs*) but the functional properties, namely the reduction potential and the hemes' order of oxidation, are markedly different. In this work, we have used nuclear magnetic resonance spectroscopy to structurally characterize PpcA from *Gm* and to probe pH-linked conformational changes in the reduced and oxidized states. The structural role of the highly conserved residue Val^{13} in the PpcA family of *Gm* and *Gs* was probed by replacing it with alanine, isoleucine, serine and threonine. Replacement of Phe^{6} and Trp^{45} in PpcA from *Gm* with the correspondent amino acids in PpcA from *Gs* – leucine and methionine – has been achieved to probe their influence in the reduction potential. The obtained results suggest that the structure in the reduced and oxidized states is conserved and similar to that of PpcA from *Gs*, with localized differences in the polypeptide segments near hemes I and III. Val^{13} has been shown to be essential for the maintenance of a single heme core conformation. The substitution of Phe^{6} and Trp^{45} yielded opposite effects on the cytochrome's reduction potential values, suggesting that these residues play different roles in the modulation of this property. These observations emphasize the preponderant role of key residues in the structure of PpcA from *Gm* and in the fine-tuning of its reduction potential values.

Keywords: *Geobacter*, extracellular electron transfer, multiheme cytochrome, nuclear magnetic resonance, protein structure

Resumo

A bactéria *Geobacter metallireducens* (*Gm*) tem a capacidade de transferir eletrões para o seu exterior e reduzir aceitadores extracelulares. Assim, é utilizada na produção de corrente elétrica quando associada a elétrodos e na biorremediação de águas contaminadas com compostos tóxicos/radioativos, como Cr(VI) ou U(VI), que atuam como aceitadores de eletrões. O citocromo *c* tri-hémico PpcA de *Gm* é extremamente abundante no periplasma e assegura a correta transferência de eletrões entre os componentes associados à membrana celular interna e externa. Embora possua 80% de identidade com o citocromo PpcA de *Geobacter sulfurreducens* (*Gs*) as propriedades funcionais de ambos – nomeadamente o potencial de redução e a ordem de oxidação dos hemos – são claramente distintas. Neste trabalho, utilizou-se a espetroscopia de ressonância magnética nuclear para caracterizar estruturalmente e estudar mudanças conformacionais associadas à variação do pH no citocromo PpcA de *Gm* no estado reduzido e oxidado. Avaliou-se o papel estrutural do resíduo Val^{13}, conservado na família PpcA de *Gm* e *Gs*, através da sua substituição por uma alanina, isoleucina, serina ou treonina. Também se substituíram os resíduos Phe^{6} e Trp^{45} pelos equivalentes no PpcA de *Gs* – leucina e metionina – e avaliou-se o impacto destas substituições no potencial de redução. Verificou-se que a estrutura do PpcA de *Gm* não se altera significativamente consoante o estado de oxidação e que é semelhante à do PpcA de *Gs* apresentando, contudo, diferenças junto dos hemos I e III. Concluiu-se que a Val^{13} é essencial para manter uma conformação única dos três hemos em solução e que os aminoácidos Phe^{6} e Trp^{45} desempenham papéis opostos na modulação do potencial de redução. Os resultados obtidos ilustram a importância de certos resíduos para a estrutura e modulação das propriedades funcionais do PpcA de *Gm*.

Palavras-chave: *Geobacter*, transferência eletrónica extracelular, citocromo multi-hémico, ressonância magnética nuclear, estrutura de proteínas

Table of Contents

List of Figures

1 Introduction

2 Materials and Methods

3 Structural and Functional Studies of PpcA in the Reduced State

4 Structural and Functional Studies of PpcA in the Oxidized State

5 Role of Phe^{6}, Val^{13} and Trp^{45} Residues in PpcA

A Appendix

List of Tables

Abbreviations, Symbols and Constants

1D	One dimensional
2D	Two dimensional
2xYT	2x yeast extract – tryptone
3D	Three dimensional
AMP	Ampicillin
ATP	Adenosine triphosphate
B_0	Magnetic Field
bp	Base pair
BLAST	Basic Local Alignment Search Tool
BMRB	Biological Magnetic Resonance Bank
Ccm	Cytochrome *c* maturation
CLO	Chloramphenicol
E. coli	*Escherichia coli*
E_{app}/E^0	Macroscopic apparent midpoint reduction potential
Da	Dalton
DIET	Direct Interspecies Electron Transfer
DMSO	Dimethyl sulfoxide
EDTA	Ethylenediaminetetraacetic acid
FID	Free Induction Decay
Gm	*Geobacter metallireducens*
Gs/GSU	*Geobacter sulfurreducens*
H-bond	Hydrogen bond
HSQC	Heteronuclear Single Quantum Coherence
I	Spin quantum number
Imc	Inner membrane cytochrome
IPTG	Isopropyl β-D-thiogalactoside
IUPAC-IUB	International Union of Pure and Applied Chemistry – International Union of Biochemistry
J	Coupling constant
LB	Luria Bertrani
lov	Lower limits for volumes
Mac	Membrane associated cytochrome
m_I	Magnetic quantum number
MFC	Microbial Fuel Cell
MQH_2	Menaquinol
MQ	Menaquinone
NADH	Nicotinamide adenine dinucleotide (reduced form)
NHE	Normal Hydrogen Electrode
NMR	Nuclear Magnetic Resonance
NOE	Nuclear Overhauser Effect
NOESY	Nuclear Overhauser Effect SpectroscopY
OD_{600}	Optical Density at 600 nm

Oma	Outer-membrane associated
Omb	Outer-membrane barrel
Omc	Outer membrane cytochrome
P	Angular momentum
P_X	Oxidation fraction *x*
p_{ijk}	Microstate with hemes *i*, *j*, *k* oxidized
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
pI	Isoelectric point
Ppc	Periplasmic cytochrome
ppm	Parts per million
RMSD	Root mean square deviation
rpm	Rotations per minute
S_X	Oxidation stage *x*
SDS-PAGE	Sodium Dodecyl Sulfate–PolyAcrylamide Gel Electrophoresis
TAE	Tris-Acetate EDTA
TEMED	Tetramethylethylenediamine
TOCSY	Total Correlation SpectroscopY
Tris	Tris(hydroxymethyl)aminomethane
upv	Upper limits for volumes
UV-Visible	Ultraviolet-Visible
WT	Wild-Type
δ	Chemical Shift
γ	Magnetogyric ratio
ε	Molar extinction coefficient
μ	Magnetic moment
t	Evolution time
τ_m	Mixing period
υ_0	Larmor frequency
$\delta_{average}$	Delta average
δ_{comb}	Combined chemical shift
F	Faraday constant (96 485 C mol^{-1})
h	Planck constant (6.63x10^{-34} m^2 kg s^{-1})
k_B	Boltzmann constant (1.38x10^{-23} m^2 kg s^{-2} K^{-1})
R	Molar gas constant (8.314 J K^{-1} mol^{-1})

Amino Acid Abbreviations

Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartate	Asp	D
Cysteine	Cys	C
Glutamate	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

Chapter 1
Introduction

1 Introduction

Anaerobic environments are prolific in microbial communities that have developed different strategies to obtain energy without using oxygen as terminal electron acceptor in the respiration processes. Dissimilatory metal-reducing organisms, such as the ones belonging to the *Geobacter* genera, couple cellular growth to the reduction of metals present in the environment [1]. This is only possible due to the development of an extracellular electron transport chain that is able to send the electrons from the cell's interior to the environment, crossing the cellular membrane. Not only this metabolism has its specificities in order to cope with electron transfer across a lipidic barrier, but also influences the biogeochemistry of the microbe's environment. Taking advantage of *Geobacter* physiological hallmarks, these bacteria have been used for bioremediation of contaminated waters and energy production in Microbial Fuel Cells (MFC).

1.1 The *Geobacter* genus

The *Geobacteraceae* family belongs to the *Desulfuromonadales* order in the *δ*-proteobacteria class and comprises four genera of Gram-negative, nonspore-forming, dissimilative iron-reducing bacteria: *Geobacter*, *Desulfuromonas*, *Desulfuromusa* and *Pelobacter* [1, 2].

Geobacter species typically colonize aquatic sediments, wetlands, rice paddies, and subsurface environments [3] where the concentration of Fe(III) in sediments frequently exceeds that of other electron acceptors such as oxygen, nitrate and sulfate [4, 5]. Not only this reduction is highly advantageous for the bacteria, but also plays an essential role in the ecosystem's geochemistry because it releases dissolved Fe(II) and Mn(II) as well as trace metals, metalloids, and phosphate that adsorb onto Fe(III) and Mn(IV) oxides. In fact, the studies that led to the discovery of the first *Geobacter* species were initially designed to better understand the flux of phosphate from aquatic sediments that contributes to algal blooms [3].

Currently, there are 19 known species of *Geobacter* isolated from different environments such as aquatic sediments (*Geobacter metallireducens*), contaminated ditches (*Geobacter sulfurreducens*), Fe(III)-reducing subsurface sediments (*Geobacter bemidjiensis*), freshwater sediments (*Geobacter lovleyi*) or uranium-contaminated subsurface sediments (*Geobacter uraniireducens*). *Geobacter* species have attracted much attention since they present a few physiological features that had never been found in living organisms: (i) use of Fe(III), Mn(IV) or humic substances as electron terminal acceptors in respiration processes; (ii) production of extracellular magnetite from microbial Fe(III) reduction; (iii) anaerobic oxidation of aromatic hydrocarbons in pure culture; (iv) microbial reduction of water contaminants such as U(VI), Cr(VI) or Co(III); (v) oxidation of organic compounds to carbon dioxide with an electrode serving as a sole electron acceptor; (vi) potential for interaction with syntrophic partners via Direct Interspecies Electron Transfer (DIET) [3].

1.2 *Geobacter metallireducens* and *Geobacter sulfurreducens*

Geobacter metallireducens (*Gm*) was the first member of the *Geobacter* species to be isolated in culture from sediments from the Potomac River Estuary, Maryland, in 1987 [5, 6]. Even though *Gm* was the first member of *Geobacter* to be described, only in 2009 its genome was fully sequenced, which allowed its genetic manipulation and study of electron transfer processes [7]. *Geobacter sulfurreducens* (*Gs*) was isolated in 1994 from surface sediments of a hydrocarbon-contaminated ditch near Norman, Oklahoma [8] and, contrarily to *Gm*, the development of a genetic system that permitted the manipulation of its genome in 2001 has established this species as the representative of the *Geobacter* genus [9].

Gm presents the unique ability of oxidizing an enormous variety of organic compounds both aromatic (such as benzaldehyde, phenol, benzene and toluene) and non-aromatic (such as butanol, ethanol and propanol) compared to *Gs,* which is only able to oxidize acetate, formate, lactate or pyruvate (only with hydrogen as electron donor) [3, 7]. Regarding the electron acceptors, both species are capable of reducing U(VI) and Tc(VII), but only *Gm* is capable of reducing V(V). This physiological hallmark has been used for the bioremediation of waters contaminated with these compounds [10-12]. The first DIET mechanism was documented in co-cultures of *Gm* and *Gs* grown in a medium with ethanol as the electron donor and fumarate as the electron acceptor [13]. DIET was required since *Gm* could metabolize ethanol but could not use fumarate as an electron acceptor and *Gs* could reduce fumarate but was not capable of metabolizing ethanol. *Gm* is also capable of performing DIET with *Methanosaeta* species in the conversion of waste waters' organic compounds to methane, an important process used for bioenergy production [14].

Besides the referred applications of *Geobacter* in the bioremediation of contaminated waters, the bacterial metabolism can be used for production of electrical current from a wide range of organic wastes in MFC devices, a promising form of sustainable energy. However, the present MFC are still far from generating a great amount of electric current [15], and several strategies are being employed for their improvement such as the design of *Geobacter* strains with enhanced extracellular electron transfer rates [16, 17]. For this, it is necessary to elucidate these mechanisms which are still poorly understood [3].

1.3 The Role of Cytochromes *c* in Extracellular Electron Transfer Mechanisms

The sequenced genomes of the *Geobacter* species have presented more than 100 genes encoding for *c*-type cytochromes with an average of 85% with more than one heme group [18]. The presence of multiple heme groups may confer several advantages: (i) the overlapping of the collective contribution of the individual heme reduction potential ranges extends the protein's global working reduction potential range; (ii) the presence of more than one heme group endows the protein with the capability of transferring more than one electron at a time, a process which is modulated by redox interactions between the hemes and/or by the existence of

protonable centers in the protein that modulate the potential according to the solution's pH (redox-Bohr effect); (iii) the capability of acting as capacitor to maintain the bacteria's metabolism during periods of time in which the cells are unable to reduce extracellular electron transfer acceptors [19].

In *Geobacter* the cytochromes are distributed along the inner membrane, periplasm and outer membrane and are responsible for guiding the electron flow from the cytoplasm to the cell's exterior, constituting an excellent target for the improvement of extracellular electron transfer rates [18].

1.3.1 Characteristics of *c*-type Cytochromes

The *c*-type cytochromes are metalloproteins with an essential role in aerobic and anaerobic respiratory processes. The heme group is constituted by a protoporphyrin IX ring with an iron atom in the center, coordinated by the four nitrogen atoms of the ring. The iron (0) is a transition metal with 26 electrons arranged in the electronic configuration $1s^2 2s^2 2p^6 3s^2 3p^6 3d^6 4s^2$, where the numbers correspond to the orbitals (s, p and d) and the superscript represents the electron occupancy. An alternative description of the electronic distribution is given by $[Ar]3d^6 4s^2$ which emphasizes the electron occupancy of the higher energy orbitals. In electron transfer cytochromes, iron generally exists in the ferrous (Fe(II)) and in the ferric (Fe(III)) states with the respective electronic configurations $[Ar]3d^6 4s^0$ and $[Ar]3d^5 4s^0$ [20]. The high and low-spin states of the iron are consequence of the distribution of the electrons in the five 3d orbitals (two orbitals e_g with higher energy and three orbitals t_{2g} with lower energy, in accordance with the Crystal Field Theory [21]), which in turn are modulated by the nature of the ligands coordinated to the atom. If the d-orbital splitting is smaller than the pairing energy (weak crystal field) the electrons tend to remain mostly unpaired in the orbitals and give rise to a high-spin state. Otherwise, if the d-orbital splitting is larger than the pairing energy (strong crystal field) the electrons tend to pair firstly on the t_{2g} orbitals and give rise to a low-spin state [20].

The heme cofactor is covalently linked to the polypeptide chain via two thioether bonds between the thiol groups of cysteines and the vinyl groups of the heme. As a consequence of this, there is a conserved binding motif sequence of the heme groups: CXXCH, where C represents the amino acid cysteine, X any given amino acid and H the amino acid histidine, which is the proximal ligand (Figure 1.1). The iron atom is always coordinated with five atoms: four nitrogen atoms of the protoporphyrin IX ring and one nitrogen atom of the imidazole ring of the histidine side chain. The sixth coordination position – that corresponds to the distal ligand – can either be transiently vacant or occupied. In cytochromes with enzymatic activity, the sixth coordination is transiently vacant [22, 23] whereas in electron transfer cytochromes it can be occupied by the side chain of a: (i) methionine, which predominates in monoheme cytochromes *c*, (ii) histidine, particularly in multiheme cytochromes or (iii) asparagine or tyrosine, which occur less frequently. In each case, the axial ligating atom is different: (i) nitrogen, in the case of histidine

or asparagine, (ii) sulfur in the case of methionine, and (iii) oxygen in the case of tyrosine [20, 22, 23].

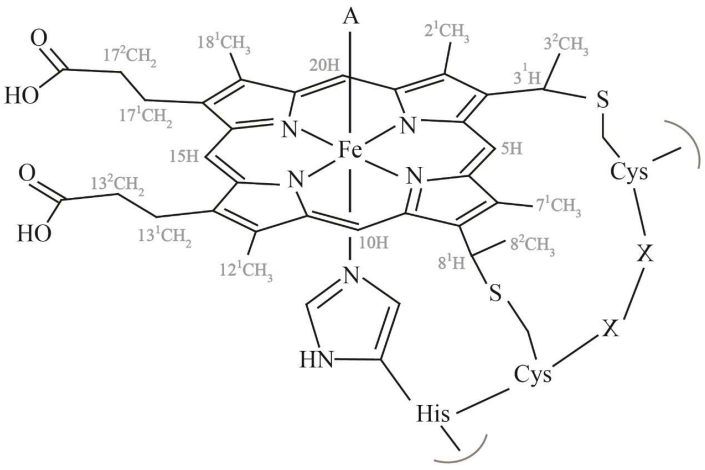


Figure 1.1 - Schematic representation of a *c*-type heme and the correspondent polypeptide binding motif. The axial coordination position labeled with A can be free or occupied by the side chain of a, for example, methionine, histidine, asparagine or tyrosine residue. The IUPAC-IUB nomenclature for tetrapyrroles is illustrated in gray [24]. Figure reproduced from [19].

1.3.2 The Extracellular Electron Transfer Chain in *Geobacter*

Currently, the extracellular electron transfer mechanisms of *Gs* have been the most studied given the earlier development of a genetic system for proteomic and gene-knockout studies in this microorganism. The current model for electron transfer (Figure 1.2) postulates that electrons originating from the oxidation of organic compounds are transferred to a menaquinone pool via a NADH hydrogenase located in the inner membrane [19]. Depending on the redox potential of the final electron acceptor, the electrons are transferred to either one of two pathways: the CbcL dependent pathway, which operates when the final electron acceptors are at or below redox potentials of -100 mV (*vs.* the normal hydrogen electrode, NHE) while the ImcH pathway is used when final electron acceptors are above the referred redox potential [25, 26]. The electrons are then transferred from CbcL or ImcH to cytochromes in the periplasm (mainly to the periplasmic cytochrome family PpcA) and from these to porin-cytochrome trans-outer membrane complexes (such as the OmaB-OmbB-OmcB or OmaC-OmbC-OmcC) which convey the electrons to the extracellular electron acceptors [27]. In addition to these electron transfer proteins, *Geobacter* bacteria possess electrically conductive *pili* that establish electrical contacts with electrodes and other methanogenic species, and flagella that allow them to move toward solid extracellular electron acceptors [28]. In this year, it was discovered that one of the extracellular cytochromes – OmcS (Outer membrane cytochrome S) –, is capable of polymerizing and forming electrically conductive filaments, illustrating the vast complexity of these mechanisms [29].

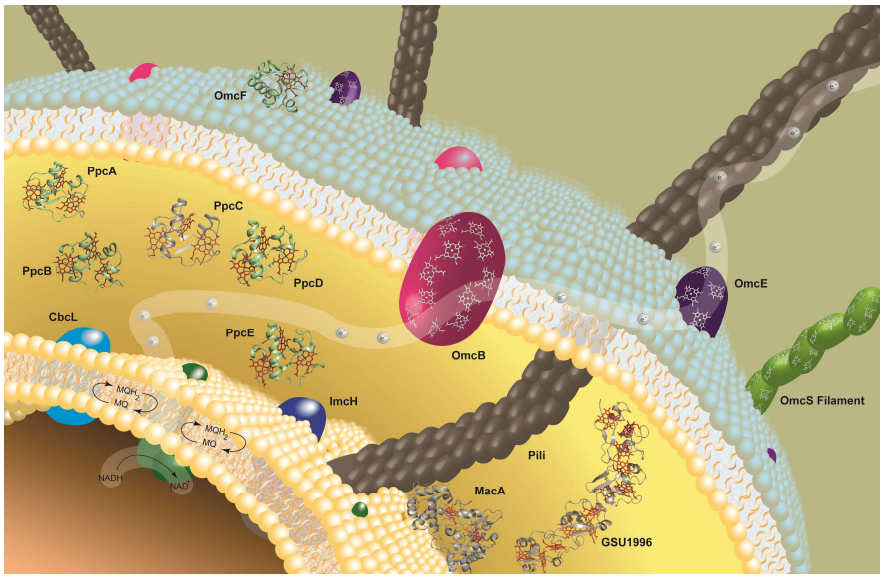


Figure 1.2 - Model of the extracellular electron transfer pathway in *Geobacter sulfurreducens*. The proposed electron transfer pathway is represented in white: the electrons from the oxidation of organic compounds are transferred to the menaquinol (MQH_2)/menaquinone (MQ) pool at the inner membrane and from here either to CbcL or ImcH cytochromes. Then, the electrons are transferred to the PpcA family of periplasmic cytochromes (PpcA (PDB code 1OS6), PpcB (PDB code 3BXU), PpcC (PDB code 3H33), PpcD (PDB code 3H4N) and PpcE (PDB code 3H34)) which mediate the electron transfer to outer membrane cytochromes. *Pili* and the newly discovered polymeric assembly of cytochrome OmcS are also represented and are thought to be involved in long-range electron transfer. MacA (PDB code 4AAL), a membrane-associated diheme cytochrome *c* which is important for reduction of Fe(III) and U(VI) oxides [30], GSU1996 (PDB code 3OV0), a "nanowire" that may function as a capacitor to enhance the bacterial electron-storage capacity [31], and OmcF (PDB code 3CU4), a monoheme cytochrome which absence negatively affects the Fe(III) citrate reduction and production of electric current [32], are also represented.

In *Gm*, the extracellular electron transfer mechanisms have been considerably less studied. The existent studies point to the conservation of some electron transfer components, such as the PpcA family of triheme cytochromes and outer membrane cytochromes OmcB and OmcE, but further structural and functional information on these components is scarce [33-36].

1.4 The PpcA Family of Triheme Cytochromes

The PpcA family of c_7 triheme cytochromes belongs to class III cytochromes since these proteins possess multiple covalently attached heme groups with low reduction potentials and bis-histidinyl coordination [20]. The PpcA family is one of the most conserved among *Geobacter*, with multiple homologs in most of the genomes: five in *G. sulfurreducens*, five in *G. metallireducens*, four in *G. uraniireducens*, three in *G. bemidjiensi*, and one in *G. lovleyi*. The PpcA family in *Gs* is constituted by PpcA, PpcB, PpcC, PpcD, PpcE whereas in *Gm* the family is constituted by PpcA, PpcB, PpcC, PpcE, PpcF. The PpcD degree of homology is very low between the *Gs* and *Gm* families and therefore the designation PpcF was adopted in *Gm* [33].

In *Gs*, PpcA, PpcB and PpcC were detected in cultures grown in the presence of Fe(III) citrate and Fe(III) oxide and the deletion of these genes affected the reduction of Fe(III) and U(VI). PpcD is more abundant during growth with Fe(III) oxides and PpcE was only detected in cultures in presence of Fe(III) citrate. The deletion of the genes of PpcD and PpcE imparted the reduction of U(VI) [37]. In *Gm*, the PpcA family is known to be essential for Fe(III) and U(VI) reduction [38] and the *ppcA* gene is generally up-regulated in DIET processes of co-cultures of *Gm* with other organisms such as *Gs* or *Methanosarcina barkeri* [39, 40].

The cytochromes from both families possess about 70 amino acids with a molecular mass of approximately 10 kDa and a pI of 9 [33, 34, 41]. The sequence alignment, as well as the identity relatively to PpcA of *Gm*, are indicated in Figure 1.3.

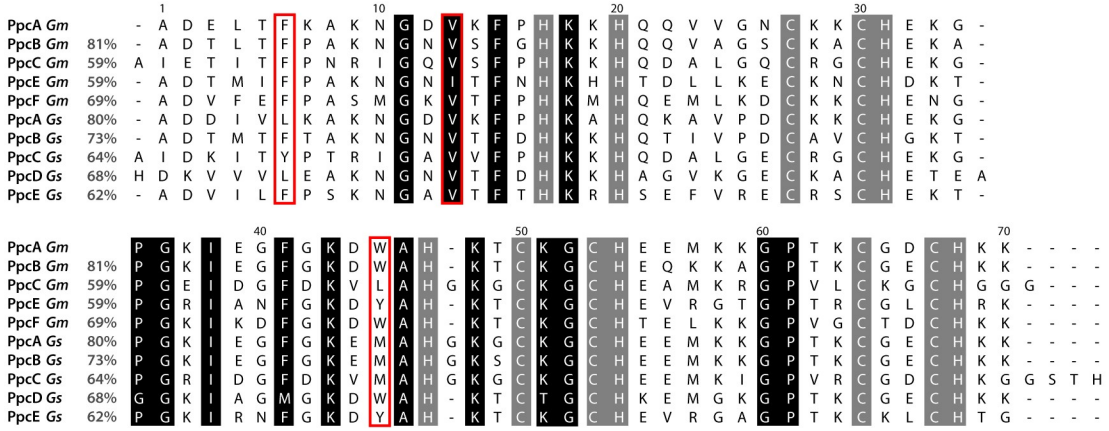


Figure 1.3 - Alignment of the amino acid sequences of the periplasmic triheme cytochromes from *Geobacter metallireducens* and *Geobacter sulfurreducens*. Black and gray boxes color the conserved non-heme attached and heme-attached residues, respectively. The residues studied within the scope of this work are highlighted in red. The heme numbering and the respective attached residues are indicated at the bottom of the last cytochrome amino acid sequence. The percentage of sequence identity relative to PpcA from *Gm* obtained from the basic local alignment search tool (BLAST) [42] is indicated. The alignment was performed with CLUSTAL Ω [43].

Three *c*-type heme groups are covalently attached to the polypeptide chain as confirmed by the existence of three CXXCH motifs in each sequence. The hemes are axially coordinated by two histidine residues and are low-spin both in the reduced diamagnetic state (Fe(II), S=0) and in the oxidized paramagnetic state (Fe(III), S=1/2) which gives rise to specific features in the UV-Visible spectrum (Figure 1.4).

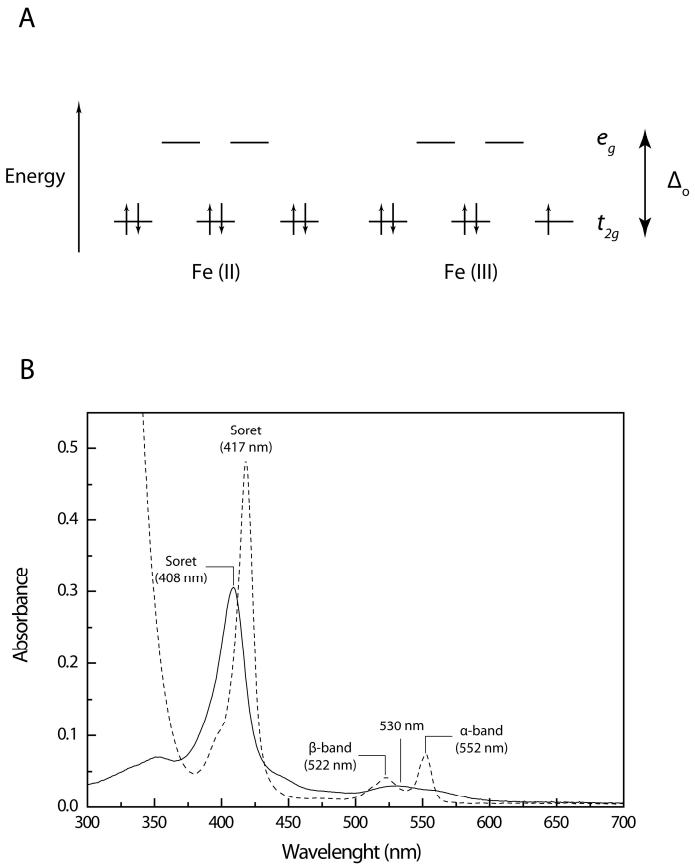


Figure 1.4 - Electronic properties of a low-spin iron attached to the heme group and UV-Visible spectroscopic properties of low-spin *c*-type cytochromes. (A) Electronic distribution of octahedral Fe(II) and Fe(III) in the low-spin state in which the d-orbital splitting, Δ_o, is higher than the energy required to pair electrons in the same orbital. (B) The spectral features are illustrated by the UV-Visible spectra of PpcA from *Gm* in the fully oxidized (solid line) and fully reduced (dashed line) states. The characteristic bands of the spectra in both redox states are indicated.

1.4.1 Structural Features of Triheme Cytochromes

To date, only the structure of the triheme cytochromes from *Gs* has been determined (Figure 1.5), showing several conserved motifs: a two-strand β-sheet segment at the N terminal, followed by several α-helixes located in different regions of the protein [44]. The spatial arrangement of the hemes in these cytochromes is conserved, with heme III nearly perpendicular to both hemes I and IV [44]. The hemes' designation in c_7 triheme cytochromes derives from the superimposition of the hemes with those of the structurally homologous tetraheme cytochromes c_3 in which the formers lack heme II and the correspondent polypeptide segment [45].

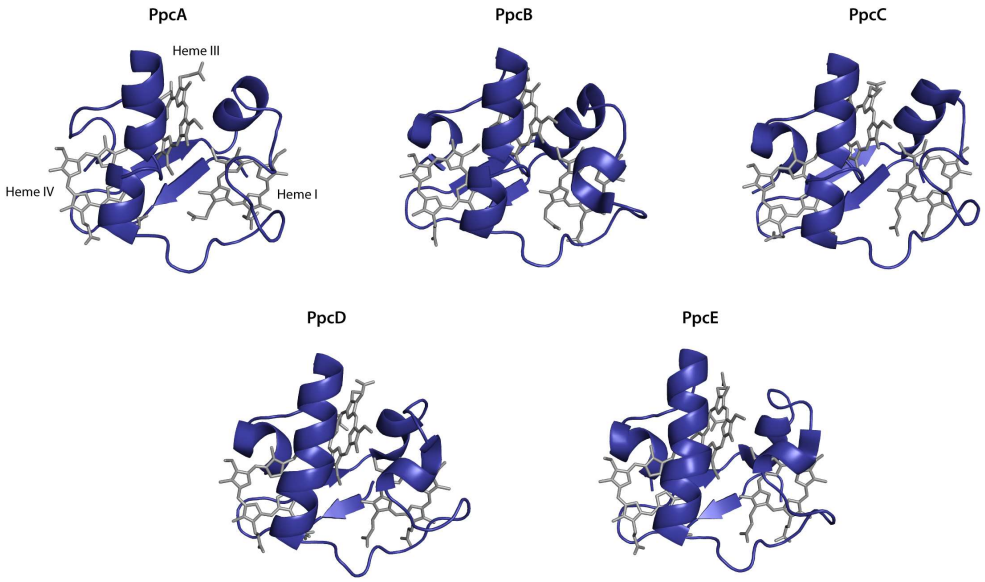


Figure 1.5 - Structures of PpcA family members from *Geobacter sulfurreducens*. PpcA is represented by its solution structure in the reduced state (PDB code 2LDO) and PpcB (PDB code 3BXU), PpcC (PDB code 3H33), PpcD (PDB code 3H4N), PpcE (PDB code 3H34) are represented by their crystal structures. PpcB and PpcD display two molecules in the crystal asymmetric unit (A and B monomers) of which monomer A is represented. The molecules are all in the same orientation. Figure produced with PyMOL [46].

Regarding the cytochromes of the PpcA family from *Gm*, only cytochromes PpcA and PpcF have been preliminarily characterized. Nuclear Magnetic Resonance (NMR) studies of these proteins in the reduced state have revealed conservation of the heme core of the proteins in comparison with PpcA from *Gs*. The assignment of the 2D $^{1}H,^{15}N$-HSQC (Heteronuclear Single Quantum Coherence) NMR spectrum of PpcA and its comparison with PpcA from *Gs* has suggested that the overall folding is maintained in both structures, with important structural differences in segments that include the cleft between hemes I and III [34]. Furthermore, the hemes' magnetic properties and axial ligands' orientation have been determined for PpcA from *Gm* in the oxidized state and have shown that there is conservation of the orientation compared to PpcA from *Gs* [47].

1.4.2 Functional Characterization of Triheme Cytochromes

In monoheme cytochromes, the redox potential can be obtained directly from the Nernst equation and the E_{app} (*i.e.* the point at which the oxidized and reduced fractions of the protein are equal) corresponds to the reduction potential of the heme. In the case of multiheme cytochromes, the co-existence of several heme groups that generally possess similar reduction potentials does not allow the direct discrimination of the individual potentials by voltammetry or potentiometric redox titrations. Instead, the E_{app} determined by these methods is denominated macroscopic reduction potential and only describes the overall redox behavior of the multiheme cytochrome.

In a triheme cytochrome, four different redox stages can be defined as a consequence of the three consecutive reversible steps of one-electron transfer that convert the fully reduced state (Stage 0, S_0) in the fully oxidized state (Stage 3, S_3) (Figure 1.6A). At each stage, several microstates with the same number of oxidized hemes co-exist and, additionally, for each microstate, the protonable center(s) in the protein that modulate the hemes' reduction potential (redox-Bohr effect) can be protonated or deprotonated, leading to a minimum of 16 possible microstates. Due to the close spatial disposition of the heme groups within the protein, the reduction potential of one heme is modulated by the oxidation stage of its neighbor (redox interactions) and by the solution pH (redox-Bohr interactions). Therefore, the complete characterization of the redox centers of a multiheme cytochrome comprises the determination of the individual heme reduction potentials, the redox interactions and the properties of the redox-Bohr center(s). Turner and co-workers have developed a strategy that permits the determination of these parameters by combining the information from UV-Visible redox titrations with NMR to determine the individual heme oxidation profiles [48]. This is possible due to the spectral signatures of the NMR spectra of low-spin multiheme cytochromes in the reduced and oxidized states (Figure 1.6B).

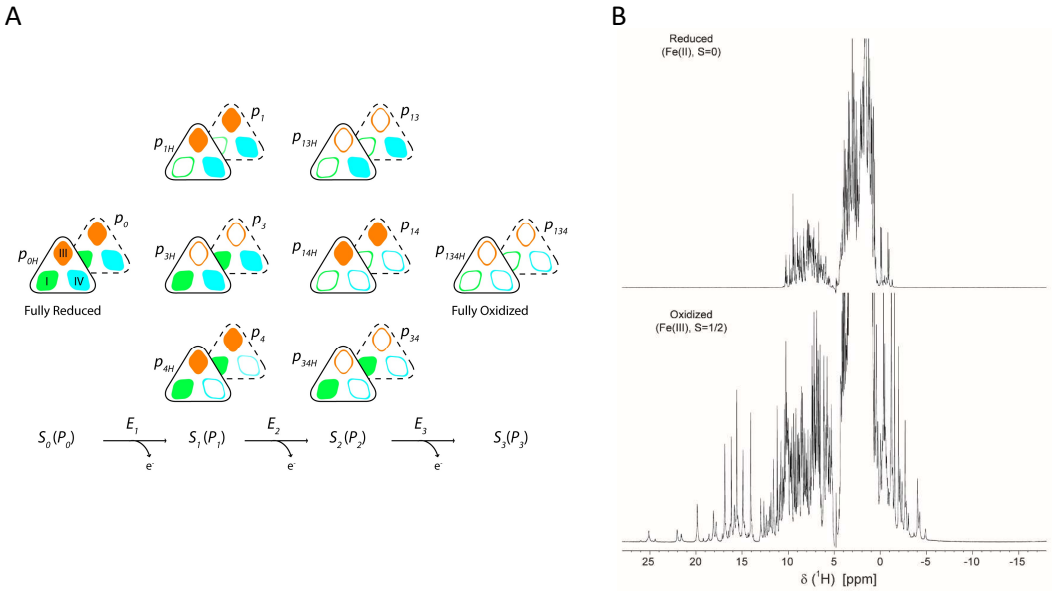


Figure 1.6 - Electronic distribution network for a triheme cytochrome and the respective 1D ^{1}H NMR spectra in the reduced and oxidized states. (A) Electronic distribution scheme for a triheme cytochrome with one proton-linked equilibrium showing the 16 possible microstates. The full and dashed triangles correspond to the protonated and deprotonated microstates, respectively. In each microstate, the heme groups are colored green (heme I), orange (heme III) and blue (heme IV), which can be either reduced (filled symbols) or oxidized (open symbols). The microstates are grouped according to the number of oxidized hemes in four oxidation stages connected by three one-electron redox steps indicated by S_0, S_1, S_2 and S_3. P_0, P_1, P_2 and P_3 indicate the molar fractions associated to each stage concerning microstates with zero, one, two or three oxidized hemes respectively. $E_{x(x\ =\ 1\text{-}3)}$ represents the macroscopic reduction potential associated with each stage. p_0 and p_{0H} indicate the deprotonated and protonated reduced microstates, respectively. p_{ijk} and p_{ijkH} correspond to the deprotonated and protonated microstates, respectively, with heme(s) *i*, *j*, and *k* oxidized. (B) 1D ^{1}H NMR spectra of the reduced (upper) and oxidized (lower) forms of PpcA from *Gm* (25 °C and pH 7). Figure reproduced from [34].

When the interconversion between microstates within the same oxidation stage – intramolecular electron exchange – is fast on the NMR timescale and the interconversion between microstates of different oxidation stages – intermolecular electron exchange – is slow, the individual heme signals in the different oxidation stages can be discriminated. Moreover, the NMR paramagnetic shifts are proportional to the degree of oxidation of a particular heme group and, therefore, can be used to monitor the oxidation of each heme. The heme methyl resonances are the easiest to identify in the NMR spectra and their largest paramagnetic shifts make them the ideal candidates for following the stepwise oxidation of the hemes throughout the redox titration. However, this information only indicates the relative heme reduction potentials and heme redox interactions. Therefore, to determine the absolute potentials, the total reduced protein fractions need to be measured through redox titrations followed by UV-Visible spectroscopy.

The known thermodynamic parameters of the PpcA family from *Gs* as well as of PpcA from *Gm* [35, 41] are summarized in Table 1.1.

Table 1.1 - Heme reduction potentials and pairwise interactions (mV) of the fully reduced and protonated forms of PpcA, PpcB, PpcD, PpcE of *Gs* and PpcA of *Gm* at 15 ºC, 250 mM ionic strength and pH 7. Reduction potentials are relative to NHE.

Protein	Heme reduction potentials			Heme-heme redox interactions			Redox-Bohr interactions		
	I	III	IV	I-III	I-IV	III-IV	I-H	III-H	IV-H
PpcA *Gs*	-154	-138	-125	27	16	41	-32	-31	-58
PpcB *Gs*	-150	-166	-125	17	8	32	-16	-9	-38
PpcD *Gs*	-156	-139	-149	46	3	14	-28	-23	-53
PpcE *Gs*	-167	-175	-116	27	5	22	-12	2	-13
PpcA *Gm*	-80	-70	-113	35	3	37	-2	-23	-49

The heme reduction potentials are negative, differ from each other and cover different functional ranges. In the particular case of PpcA from *Gm* and PpcA from *Gs*, the thermodynamic properties are markedly different despite the 80% of sequence identity. Furthermore, at pH 7, the heme oxidation order is IV-I-III for PpcA from *Gm* and I-IV-III for PpcA from *Gs*, with these proteins displaying different preferential pathways for electron transfer (Figure 1.7) but maintaining the ability to couple proton to electron transfer. Therefore, it is striking that two proteins with such similarity in the amino acid content present markedly different functional properties.

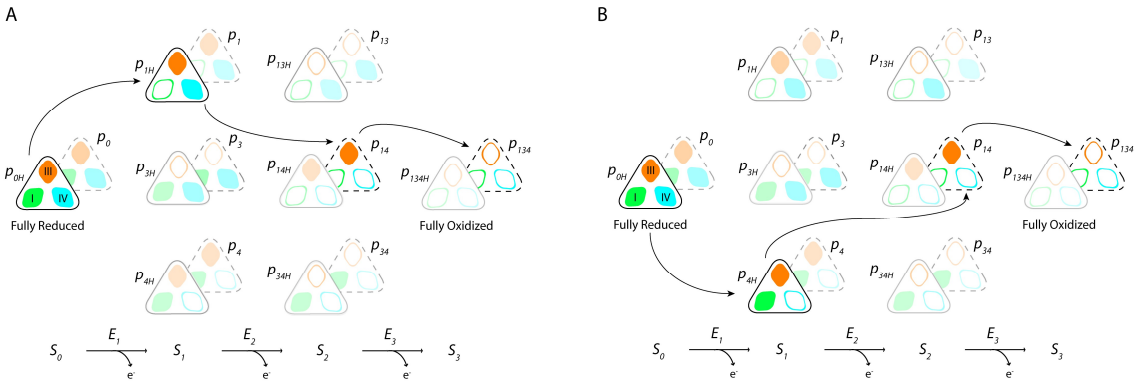


Figure 1.7 - Preferential pathway for electron/proton transfer in PpcA from (A) *Geobacter sulfurreducens* and (B) *Geobacter metallireducens* at pH 7. The black arrows indicate the preferential electron transfer pathway involving the relevant microstates. An electron/proton coupling is observed between oxidation stages S_1 and S_2.

1.5 Objectives and Thesis Outline

The functional differences shown by PpcA from *Gm* and PpcA from *Gs*, regardless their high sequence similarity, establish these two proteins as an excellent model to identify the key residues that account for the observed functional differences. However, to attain this objective, it is first necessary to establish a solid characterization of PpcA from *Gm* in both the reduced and oxidized states.

In this thesis, NMR spectroscopy was used to structurally characterize PpcA from *Gm* in the reduced and oxidized states and to probe the structural and functional impact of residues Val[13], Phe[6] and Trp[45] in the modulation of redox properties. Chapter 1 provides a general contextualization of *Geobacter* and the role of multiheme cytochromes in extracellular electron transfer mechanisms. Chapter 2 describes the experimental procedures employed in this work and introduces the fundamental principles of NMR spectroscopy. Chapters 3 and 4 describe the structural and functional characterization of PpcA in the reduced and oxidized states, respectively. Site-directed mutagenesis on the above mentioned residues was carried out and the structural and functional characterization of these mutants are described in Chapter 5. Finally, Chapter 6 highlights the main conclusions of this thesis and presents future directions that can be pursued.

1.6 References

[1] M. Madigan, J. Martinko, K. Bender, D. Buckley, D. Stahl, Brock Biology of Microorganisms, 14 ed., Benjamin Cummings, Boston, 2014.
[2] W.F.M. Röling, The family *Geobacteraceae*, in: E. Rosenberg, E.F. DeLong, S. Lory, E. Stackebrandt, F. Thompson (Eds.) The Prokaryotes: *Deltaproteobacteria* and *Epsilonproteobacteria*, Springer Berlin Heidelberg, Berlin, Heidelberg, 2014, pp. 157-172.
[3] D.R. Lovley, T. Ueki, T. Zhang, N.S. Malvankar, P.M. Shrestha, K.A. Flanagan, M. Aklujkar, J.E. Butler, L. Giloteaux, A.-E. Rotaru, D.E. Holmes, A.E. Franks, R. Orellana, C. Risso, K.P. Nevin, *Geobacter*: The Microbe Electric's Physiology, Ecology, and Practical Applications, in: R.K. Poole (Ed.) Advances in Microbial Physiology, Academic Press, 2011, pp. 1-100.
[4] D.R. Lovely, Microbial reduction of iron, manganese, and other metals, Advances in Agronomy (USA), v. 54 (1995).
[5] D.R. Lovley, E.J. Phillips, Organic matter mineralization with reduction of ferric iron in anaerobic sediments, Applied and Environmental Microbiology, 51 (1986) 683-689.
[6] D.R. Lovley, E.J. Phillips, Novel mode of microbial energy metabolism: organic carbon oxidation coupled to dissimilatory reduction of iron or manganese, Applied and Environmental Microbiology, 54 (1988) 1472-1480.
[7] M. Aklujkar, J. Krushkal, G. DiBartolo, A. Lapidus, M.L. Land, D.R. Lovley, The genome sequence of *Geobacter metallireducens*: features of metabolism, physiology and regulation common and dissimilar to *Geobacter sulfurreducens*, BMC Microbiology, 9 (2009) 109.
[8] F. Caccavo, Jr., D.J. Lonergan, D.R. Lovley, M. Davis, J.F. Stolz, M.J. McInerney, *Geobacter sulfurreducens* sp. nov., a hydrogen- and acetate-oxidizing dissimilatory metal-reducing microorganism, Applied and Environmental Microbiology, 60 (1994) 3752-3759.
[9] M.V. Coppi, C. Leang, S.J. Sandler, D.R. Lovley, Development of a genetic system for *Geobacter sulfurreducens*, Applied and Environmental Microbiology, 67 (2001) 3180-3187.
[10] G. Liu, J. Zhou, C. Chen, J. Wang, R. Jin, H. Lv, Decolorization of azo dyes by *Geobacter metallireducens*, Applied Microbiology and Biotechnology, 97 (2013) 7935-7942.
[11] I. Ortiz-Bernad, R.T. Anderson, H.A. Vrionis, D.R. Lovley, Vanadium respiration by *Geobacter metallireducens*: novel strategy for in situ removal of vanadium from groundwater, Applied and Environmental Microbiology, 70 (2004) 3091-3095.
[12] J.R. Lloyd, V.A. Sole, C.V. Van Praagh, D.R. Lovley, Direct and Fe(II)-mediated reduction of technetium by Fe(III)-reducing bacteria, Applied and Environmental Microbiology, 66 (2000) 3743-3749.
[13] Z.M. Summers, H.E. Fogarty, C. Leang, A.E. Franks, N.S. Malvankar, D.R. Lovley, Direct exchange of electrons within aggregates of an evolved syntrophic coculture of anaerobic bacteria, Science, 330 (2010) 1413-1415.
[14] A.-E. Rotaru, P.M. Shrestha, F. Liu, M. Shrestha, D. Shrestha, M. Embree, K. Zengler, C. Wardman, K.P. Nevin, D.R. Lovley, A new model for electron flow during anaerobic digestion: direct interspecies electron transfer to *Methanosaeta* for the reduction of carbon dioxide to methane, Energy & Environmental Science, 7 (2014) 408-415.
[15] A.E. Franks, K.P. Nevin, Microbial fuel cells, a current review, Energies, 3 (2010) 899-919.
[16] M. Izallalen, R. Mahadevan, A. Burgard, B. Postier, R. Didonato Jr, J. Sun, C.H. Schilling, D.R. Lovley, *Geobacter sulfurreducens* strain engineered for increased rates of respiration, Metabolic engineering, 10 (2008) 267-275.

[17] H. Yi, K.P. Nevin, B.-C. Kim, A.E. Franks, A. Klimes, L.M. Tender, D.R. Lovley, Selection of a variant of *Geobacter sulfurreducens* with enhanced capacity for current production in microbial fuel cells, Biosensors and Bioelectronics, 24 (2009) 3498-3503.
[18] J.E. Butler, N.D. Young, D.R. Lovley, Evolution of electron transfer out of the cell: comparative genomics of six *Geobacter* genomes, BMC Genomics, 11 (2010) 40.
[19] C.A. Salgueiro, J.M. Dantas, Multiheme Cytochromes, in: Multiheme Cytochromes, Springer Berlin Heidelberg, Berlin, Heidelberg, 2016, pp. 1-39.
[20] G.R. Moore, G.W. Pettigrew, Cytochromes *c*: evolutionary, structural and physicochemical aspects, Springer Science & Business Media, 2012.
[21] J. Van Vleck, Theory of the variations in paramagnetic anisotropy among different salts of the iron group, Physical Review, 41 (1932) 208.
[22] I. Bertini, G. Cavallaro, A. Rosato, Cytochrome *c*: occurrence and functions, Chemical Reviews, 106 (2006) 90-115.
[23] L.J. Smith, A. Kahraman, J.M. Thornton, Heme proteins—diversity in structural characteristics, function, and folding, Proteins: structure, function, and bioinformatics, 78 (2010) 2349-2368.
[24] G.P. Moss, Nomenclature of tetrapyrroles, European Journal of Biochemistry, 178 (1988) 277-328.
[25] L. Zacharoff, C.H. Chan, D.R. Bond, Reduction of low potential electron acceptors requires the CbcL inner membrane cytochrome of *Geobacter sulfurreducens*, Bioelectrochemistry, 107 (2016) 7-13.
[26] C.E. Levar, C.H. Chan, M.G. Mehta-Kolte, D.R. Bond, An inner membrane cytochrome required only for reduction of high redox potential extracellular electron acceptors, Mbio, 5 (2014) 1-9.
[27] Y. Liu, Z. Wang, J. Liu, C. Levar, M.J. Edwards, J.T. Babauta, D.W. Kennedy, Z. Shi, H. Beyenal, D.R. Bond, T.A. Clarke, J.N. Butt, D.J. Richardson, K.M. Rosso, J.M. Zachara, J.K. Fredrickson, L. Shi, A trans-outer membrane porin-cytochrome protein complex for extracellular electron transfer by *Geobacter sulfurreducens* PCA, Environmental Microbiology Reports, 6 (2014) 776-785.
[28] P.-L. Tremblay, M. Aklujkar, C. Leang, K.P. Nevin, D. Lovley, A genetic system for *Geobacter metallireducens*: role of the flagellin and pilin in the reduction of Fe(III) oxide, Environmental Microbiology Reports, 4 (2012) 82-88.
[29] F. Wang, Y. Gu, J.P. O'Brien, S.M. Yi, S.E. Yalcin, V. Srikanth, C. Shen, D. Vu, N.L. Ing, A.I. Hochbaum, E.H. Egelman, N.S. Malvankar, Structure of microbial nanowires reveals stacked hemes that transport electrons over micrometers, Cell, 177 (2019) 361-369.e310.
[30] B.-C. Kim, D.R. Lovley, Investigation of direct vs. indirect involvement of the *c*-type cytochrome MacA in Fe(III) reduction by *Geobacter sulfurreducens*, FEMS Microbiology Letters, 286 (2008) 39-44.
[31] P.R. Pokkuluri, Y.Y. Londer, N.E.C. Duke, M. Pessanha, X. Yang, V. Orshonsky, L. Orshonsky, J. Erickson, Y. Zagyanskiy, C.A. Salgueiro, M. Schiffer, Structure of a novel dodecaheme cytochrome *c* from *Geobacter sulfurreducens* reveals an extended 12 nm protein with interacting hemes, Journal of Structural Biology, 174 (2011) 223-233.
[32] J.M. Dantas, M.A. Silva, D. Pantoja-Uceda, D.L. Turner, M. Bruix, C.A. Salgueiro, Solution structure and dynamics of the outer membrane cytochrome OmcF from *Geobacter sulfurreducens*, Biochimica et Biophysica Acta - Bioenergetics, 1858 (2017) 733-741.

[33] M.R. Ferreira, J.M. Dantas, C.A. Salgueiro, The triheme cytochrome PpcF from *Geobacter metallireducens* exhibits distinct redox properties, FEBS Open Bio, 8 (2018) 1897-1910.
[34] P.C. Portela, T.M. Fernandes, J.M. Dantas, M.R. Ferreira, C.A. Salgueiro, Biochemical and functional insights on the triheme cytochrome PpcA from *Geobacter metallireducens*, Archives of Biochemistry and Biophysics, 644 (2018) 8-16.
[35] T.M. Fernandes, L. Morgado, C.A. Salgueiro, Thermodynamic and functional characterization of the periplasmic triheme cytochrome PpcA from *Geobacter metallireducens*, Biochemical Journal, 475 (2018) 2861-2875.
[36] J.A. Smith, D.R. Lovley, P.-L. Tremblay, Outer cell surface components essential for Fe(III) oxide reduction by *Geobacter metallireducens*, Appl. Environ. Microbiol., 79 (2013) 901-907.
[37] Y.-H.R. Ding, K.K. Hixson, M.A. Aklujkar, M.S. Lipton, R.D. Smith, D.R. Lovley, T. Mester, Proteome of *Geobacter sulfurreducens* grown with Fe(III) oxide or Fe(III) citrate as the electron acceptor, Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics, 1784 (2008) 1935-1941.
[38] J.R. Lloyd, C. Leang, A.L.H. Myerson, M.V. Coppi, S. Cuifo, B. Methe, S.J. Sandler, D.R. Lovley, Biochemical and genetic characterization of PpcA, a periplasmic *c*-type cytochrome in *Geobacter sulfurreducens*, Biochemical Journal, 369 (2003) 153-161.
[39] P.M. Shrestha, A.-E. Rotaru, Z.M. Summers, M. Shrestha, F. Liu, D.R. Lovley, Transcriptomic and Genetic Analysis of Direct Interspecies Electron Transfer, Applied and Environmental Microbiology, 79 (2013) 2397-2404.
[40] A.-E. Rotaru, P.M. Shrestha, F. Liu, B. Markovaite, S. Chen, K.P. Nevin, D.R. Lovley, Direct Interspecies Electron Transfer between *Geobacter metallireducens* and *Methanosarcina barkeri*, Applied and Environmental Microbiology, 80 (2014) 4599-4605.
[41] L. Morgado, M. Bruix, M. Pessanha, Y.Y. Londer, C.A. Salgueiro, Thermodynamic characterization of a triheme cytochrome family from *Geobacter sulfurreducens* reveals mechanistic and functional diversity, Biophysical Journal, 99 (2010) 293-301.
[42] S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D.J. Lipman, Basic local alignment search tool, Journal of Molecular Biology, 215 (1990) 403-410.
[43] F. Madeira, Y.M. Park, J. Lee, N. Buso, T. Gur, N. Madhusoodanan, P. Basutkar, A.R.N. Tivey, S.C. Potter, R.D. Finn, R. Lopez, The EMBL-EBI search and sequence analysis tools APIs in 2019, Nucleic Acids Research, 47 (2019) W636-W641.
[44] P.R. Pokkuluri, Y.Y. Londer, X. Yang, N.E.C. Duke, J. Erickson, V. Orshonsky, G. Johnson, M. Schiffer, Structural characterization of a family of cytochromes c_7 involved in Fe(III) respiration by *Geobacter sulfurreducens*, Biochimica et Biophysica Acta (BBA) - Bioenergetics, 1797 (2010) 222-232.
[45] D.L. Turner, H.S. Costa, I.B. Coutinho, J. LeGall, A.V. Xavier, Assignment of the ligand geometry and redox potentials of the trihaem ferricytochrome c_3 from *Desulfuromonas acetoxidans*, Eur. J. Biochem., 243 (1997) 474-481.
[46] Schrodinger, LLC, The PyMOL Molecular Graphics System, Version 1.8, in, 2015.
[47] T.M. Fernandes, L. Morgado, C.A. Salgueiro, D.L. Turner, Determination of the magnetic properties and orientation of the heme axial ligands of PpcA from *Geobacter metallireducens* by paramagnetic NMR, Journal of Inorganic Biochemistry, 198 (2019) 110718.
[48] D.L. Turner, C.A. Salgueiro, T. Catarino, J. LeGall, A.V. Xavier, NMR studies of cooperativity in the tetrahaem cytochrome c_3 from *Desulfovibrio vulgaris*, European Journal of Biochemistry, 241 (1996) 723-731.

Chapter 2
Materials and Methods

2 Materials and Methods

2.1 Expression Vectors and Site-directed Mutagenesis

The expression vector containing the *ppcA* gene from *Gm* – pCSGmet2902 – was already available in the host laboratory. This plasmid contains a *lac* promotor, an ompA signal sequence from *Escherichia coli* (*E. coli*) and a resistance marker to ampicillin [1].

For the site-directed mutagenesis, the protocol followed a strategy based on the NZYMutagenesis kit (NZYTech), which is indicated in Figure 2.1.

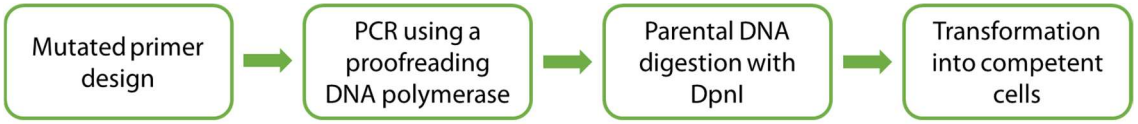


Figure 2.1 - Overview of the NZYMutagenesis kit protocol. Adapted from the kit's instructions manual.

The primers for F6L and W45M mutants were designed by the QuikChange Primer Design program (Agilent Technologies) and synthesized by Invitrogen. For the F6L mutant, the primers used were 5'-CCG CTG ACG AGC TTA CCT TAA AGG CAA AGA ACG-3' (reverse primer, 52% GC content) and 5'-CGT TCT TTG CCT TTA AGG TAA GCT CGT CAG CGG-3' (forward primer, 52% GC content) and for the W45M mutant the primers used were 5'-ATC GAG GGC TTT GGC AAG GAT ATG GCT CAC AAG ACT TG-3' (reverse primer, 50% GC content) and 5'-CAA GTC TTG TGA GCC ATA TCC TTG CCA AAG CCC TCG AT-3' (forward primer, 50% GC content). The pCSGmet2902 template was amplified by Polymerase Chain Reaction (PCR) using the primers harboring the desired mutations. Due to the high GC content of the primers, two PCR mix reactions using different buffers (Phusion High Fidelity Buffer or GC Buffer) and dimethyl sulfoxide (DMSO) were tested according to the described in Table 2.1. The PCR program is described in Table 2.2. After selection of the best conditions, the parental DNA was eliminated from the selected PCR products by digestion with DpnI (NZYTech), which is specific for methylated DNA, for 1 h at 37 °C after which the enzyme was inactivated by 20 min incubation at 80 °C. The size of the PCR products was confirmed by 0.8% agarose gel using GeneRuler 1kb DNA ladder (New England BioLabs).

Table 2.1 - PCR mix components used for site-directed mutagenesis of PpcA. DMSO is only used with GC Buffer. The Phusion Buffers, DMSO, nucleotides and Phusion DNA polymerase are all from Thermo Fisher Scientific. 1 U is the amount of enzyme needed to convert in one minute 1 µmole of substrate.

PCR mix components	Final concentration
Phusion High Fidelity Buffer or GC Buffer	1x
Nucleotides	200 µM
Forward Primer	0.5 µM
Reverse Primer	0.5 µM
Template DNA	0.6 ng/µL
Phusion DNA Polymerase	0.02 U/µL
DMSO	3 %

Table 2.2 - PCR program used for site-directed mutagenesis of PpcA.

PCR step	Temperature (ºC)	Time (min)	Cycles
Initial denaturation	98	2	1
Denaturation	98	1	30
Annealing	83	1.5	
Extension	72	6.75	
Final extension	72	15	1
Final hold	4	-	-

E. coli DH5α competent cells were transformed with the PCR products using the Heat Shock Method [2] where cells were incubated 30 minutes on ice, followed by 1-minute incubation at 42 °C and 1-minute incubation on ice again. Then, 500 µL of Luria Bertrani (LB) medium was added and cells were grown aerobically for 1 h at 37 °C, 200 rpm. After the transformation step, the cells were plated on solid LB medium supplemented with 100 µg/mL of ampicillin (AMP) (NZYTech) and incubated overnight at 37 °C. Positive and negative controls were always carried out. Colonies harboring the plasmid with resistance to AMP were inoculated into 5 mL of liquid LB medium supplemented with 100 µg/mL of AMP and were grown overnight aerobically at 37 °C, 200 rpm. The plasmid was isolated from the grown culture using the NZYMiniprep kit (NZYTech) according to the manufacturer's instructions and the presence of desired mutations was confirmed by DNA sequencing in both strands by STAB VIDA.

2.2 Protein Expression and Purification

The protein expression and purification procedure follows the described by Portela and co-workers [3] and applies to all proteins studied in this thesis. In the scope of this thesis, F6L and W45M mutants both in natural abundance and ^{15}N-isotopically labeled were expressed and

purified. Wild-type PpcA from *Gm* both in natural abundance and ^{13}C, ^{15}N-isotopically labeled were already available in the host lab, as well as PpcA V13 mutants from *Gs*.

All the descriptions given below apply both to natural abundance and ^{15}N-isotopically labeled protein overexpression protocols, except for sections 2.2.2 and 2.2.3.

2.2.1 Transformation of *E. coli* (BL21(DE3)) Cells

Competent *E. coli* BL21(DE3) cells harboring the plasmid pEC86 (that encodes for cytochrome *c* maturation gene cluster *ccmABCDEFH* and with a chloramphenicol (CLO) resistance marker) were transformed with 50 ng of plasmid containing the gene encoding to F6L or W45M mutants following the Heat Shock Method procedure [2] described in section 2.1. Then, cells were incubated with 500 µL of liquid 2x yeast extract – tryptone (2xYT) medium for 1 h at 37 °C, 200 rpm and plated onto solid 2xYT medium supplemented with 100 µg/mL of AMP and 34 µg/mL of CLO (NZYTech). The plates were incubated overnight at 37 °C. Positive and negative controls were also always carried out.

2.2.2 Natural Abundance Protein Overexpression

A colony of transformants was selected and inoculated in 50 mL of liquid 2xYT medium supplemented with 100 µg/mL of AMP, 34 µg/mL of CLO and was incubated aerobically overnight at 30 °C, 200 rpm. On the following day, 10 mL of this culture were transferred to 1 L of liquid 2xYT medium supplemented with the same concentrations of AMP and CLO and the cells were incubated aerobically at 30 °C, 180 rpm until they reached an OD_{600} between 1.5 and 1.8. Then, protein overexpression was induced with a final concentration of 100 µM of isopropyl β-D-thiogalactoside (IPTG) and cells were incubated overnight at 30 °C, 160 rpm.

2.2.3 Isotopically Labeled Protein Overexpression

A colony of transformants was inoculated in 5 mL of liquid 2xYT medium supplemented with the same concentration of AMP and CLO as in section 2.2.2 and grown aerobically overnight at 30 °C, 200 rpm. Then, 500 µL of the culture were transferred to 50 mL of 2xYT liquid medium supplemented with AMP and CLO followed by incubation at 30 °C, 200 rpm until an OD_{600} between 1.5 and 1.8 was reached. After this, 10 mL of the culture were inoculated in 1 L of 2xYT medium supplemented with AMP and CLO in the usual concentrations and cells were incubated 30 °C, 180 rpm to an OD_{600} of 1.5–1.8. Cells were then collected by centrifugation (Avanti J-26 XPI Beckman Coulter) – 6 400 *xg*, 20 min, 20 °C –, washed twice with 250 mL of salt solution 1xM9 (containing 3 g/L of KH_2PO_4 (Riedel-de-Haen), 6 g/L of Na_2HPO_4 (VWRChemicals), 0.5 g/L of NaCl (NZYTech)) and then transferred to minimal medium M9 (in a ratio of 250 mL of minimal medium for each liter of 2xYT medium) containing 2.2 mM

KH_2PO_4 (Riedel-de-Haen), 3.37 mM Na_2HPO_4 (VWRChemicals), 0.86 mM NaCl (Panreac), 20 mg/mL biotin (Fagron), 5 µM $MnCl_2 \cdot 4H_2O$ (VWRChemicals), 2 mM $MgSO_4 \cdot 7H_2O$ (Panreac), 0.1 mM $CaCl_2$ (Sigma-Aldrich), 100 µM $FeSO_4 \cdot 7H_2O$ (Merck), supplemented with 1 mM δ-aminolevulinic acid (Merck), 59.2 µM thiamine (Merck), 4 mg/L D-glucose ^{12}C (VWRChemicals), 1 mg/L $^{15}NH_4Cl$ (Cambridge Isotope Laboratories - CIL). After 90 minutes of incubation at 30 °C, 180 rpm, protein expression was induced with a final concentration of 100 µM IPTG and cells were grown overnight at 30 °C, 160 rpm.

2.2.4 Protein Isolation and Purification

Cells were harvested by centrifugation (Avanti J-26 XPI Beckman Coulter) – 4 000 *xg*, 20 min, 4 °C – and the periplasmic fraction was isolated using lysis buffer containing 200 mM Tris–HCl (NZYTech), pH 8, 0.2 mM EDTA (Amresco), 20 % sucrose (VWR Chemicals) and 0.5 mg/mL of lysozyme (Fluka). After centrifugation at 14 700 *xg*, 20 min, 4 °C, the red supernatant was ultracentrifuged (Ultracentrifuge Optima LE-80K Beckman Coulter) at 225 000 *xg*, 1 h, 4 °C and then dialyzed against 2x 10 mM Tris-HCl, pH 8 buffer.

Purification encompassed a cation exchange chromatography step where the periplasmic fraction was loaded onto a 2x5 mL Bio-Scale™ Mini UNOsphere™ S Cartridges (BioRad) equilibrated with the same dialysis buffer. The protein was eluted with a 150 mL gradient of 0-300 mM NaCl at a flow rate of 1 mL/min. The red fractions containing PpcA were concentrated to 1.5 mL and were then loaded onto a Hiload 16/60 Superdex 75 column (GE Healthcare), with a molecular weight range between 3 and 70 kDa, equilibrated with 100 mM sodium phosphate buffer, pH 8 and eluted at a flow rate of 0.5 mL/min. Both chromatographic steps were performed with an ÄKTA Prime Plus Chromatography System (GE Healthcare).

2.2.5 Protein Purity Evaluation and Quantification

The protein purity was evaluated by SDS-PAGE gel (5% acrylamide stacking gel and 15% acrylamide running gel). Samples were prepared with an equal volume of SDS-PAGE Sample Loading Buffer (NZYTech) and were separated using a Mini-PROTEAN® Electrophoresis System by applying a 120 V current for 90 minutes. The gels were stained with BlueSafe (NZYTech) and either Protein Marker Precision Plus Protein™ Dual Xtra Standards (BioRad) ranging 250-2 kDa or NZYColour Protein Marker I ranging 245-5 kDa (NZYTech) were used to verify the protein's molecular weight.

For protein quantification, UV-Visible absorption spectra in the oxidized and reduced states were acquired at room temperature on a UV–Visible scanning spectrophotometer Ultraspec 2100pro (Amersham Biosciences, Switzerland) using quartz cuvettes with 1 cm path length. Protein reduction was achieved by adding crystals of sodium dithionite (Thermo Fisher

Scientific) to the sample. Protein quantification was estimated using the molar extinction coefficient for reduced PpcA at 552 nm (ε_{552nm}=118 mM^{-1} cm^{-1}) [3].

2.3 Redox Titrations Followed by UV-Visible Spectroscopy

The redox titrations of F6L and W45M mutants were followed by UV-Visible spectroscopy inside an anaerobic LABstar glove box (MBraun) with argon circulation and oxygen levels kept under 0.5 ppm, as previously described for the wild-type protein [3]. The UV-Visible spectra were acquired with a Thermo Scientific Evolution™ 300 UV-Visible spectrophotometer and the sample temperature was maintained at 15 °C by using an external circulating bath. Solutions containing 18 μM of protein were prepared in 80 mM phosphate buffer with NaCl (250 mM final ionic strength), pH 8. The solution reduction potential value was measured using a combined Pt/Ag/AgCl electrode (Crison), calibrated at each titration with freshly prepared saturated solutions of quinhydrone (Merck) at pH 7 and 4 and checked at the end for stability. In order to ensure a good equilibrium between the redox centers and the working electrode, a mixture of the following redox mediators was added to the protein solution, all with approximately 4 μM final concentration: gallocyanine ($E^{0'}$ = +21 mV), methylene blue ($E^{0'}$ = +11 mV), indigo tetrasulfonate ($E^{0'}$ = -30 mV), indigo trisulfonate ($E^{0'}$ = -70 mV), indigo disulfonate ($E^{0'}$ = -110 mV), 2-hidroxy-1,4-naphthoquinone ($E^{0'}$ = -152 mV), anthraquinone-2,6-disulfonate ($E^{0'}$ = -184 mV), anthraquinone-2-sulfonate ($E^{0'}$ = -225 mV), safranine O ($E^{0'}$ = -280 mV), neutral red ($E^{0'}$ = -325 mV), benzyl viologen ($E^{0'}$ = -345 mV), diquat ($E^{0'}$ = -350 mV) and methyl viologen ($E^{0'}$ = -440 mV). Each titration was performed at least two times both in the oxidative and reductive directions to check for hysteresis and reproducibility using sodium dithionite (Thermo Fisher Scientific) as reducing agent and potassium ferricyanide (Merck) as oxidizing agent. The reduced fraction of PpcA was determined by integrating the area of the α-peak (552 nm) above the line connecting the flanking isosbestic points (545 and 560 nm) to subtract the optical contribution of the redox mediators. The fitting of the experimental data used Equation 1 which deduction is presented in section A3 – Appendix A.

$$\text{Reduced fraction} = \frac{3+2e^{\left[\left(E-E_1\right)\frac{F}{RT}\right]}+e^{\left[\left(2E-E_1-E_2\right)\frac{F}{RT}\right]}}{3\left(1+e^{\left[\left(E-E_1\right)\frac{F}{RT}\right]}+e^{\left[\left(2E-E_1-E_2\right)\frac{F}{RT}\right]}+e^{\left[\left(3E-E_1-E_2-E_3\right)\frac{F}{RT}\right]}\right)} \quad (1)$$

2.4 Nuclear Magnetic Resonance Studies

All NMR experiments were carried out on a Bruker Avance 600 MHz spectrometer equipped with a triple-resonance cryoprobe (TCI) at 25 °C. ^{1}H chemical shifts were calibrated using the water signal as internal reference and ^{13}C and ^{15}N chemical shifts were calibrated through indirect referencing [4]. The NMR spectra were processed using TOPSPIN software (Bruker Biospin, Karlsruhe, Germany) and analyzed with Sparky Software (TD Goddard and DG Kneller, Sparky 3, University of California, San Francisco, USA).

2.4.1 NMR Fundamentals

NMR relies on the magnetic properties of atomic nuclei to study their selective absorption of electromagnetic radiation under the presence of an applied external magnetic field.

All nuclei possess angular momentum, P, and a magnetic moment, μ. These two quantized variables are related by the magnetogyric ratio, γ, (Equation 2) which is characteristic of each nucleus.

$$\mu = \gamma P \tag{2}$$

The angular momentum P is represented by the spin quantum number, I, that can take half-integer or integer values. Only nuclei with I ≠ 0 can be detected on NMR and the total number of possible energy levels, represented by the magnetic quantum number m_I, is of 2I+1.

The most biologically relevant nuclei, such as ^{1}H, ^{13}C, ^{15}N, ^{19}F and ^{31}P, possess I = 1/2 and, therefore, present two degenerate states: α (m_I = 1/2) and β (m_I = -1/2). When a static, external, magnetic field B_0 is applied, the degeneracy of the energy levels is abolished because the nuclear magnetic moment interacts differently with B_0 and, provided the nuclei's γ is positive, state β will be more energetic than state α. Nuclei will populate α or β states according to the Boltzmann equation (Equation 3).

$$\frac{N_\alpha}{N_\beta} = e^{\frac{\Delta E}{k_B T}} \tag{3}$$

In this equation, $\frac{N_\alpha}{N_\beta}$ corresponds to the ratio between α and β populations, k_B to the Boltzmann's constant and T to the temperature. There will be an excess of nuclei in state α because it is the less energetic one, leading to a net nuclear magnetic moment commonly known as magnetization (M), represented by a vector parallel to the static magnetic field B_0.

The energy difference, ΔE, between α and β states that arises upon application of B_0 is called the nuclear Zeeman splitting. To stimulate an energy transition between both states, the energy provided must be exactly the same as ΔE, which is given by Equation 4 where h represents Planck's constant:

$$\Delta E = \frac{\gamma h B_0}{4\pi} \tag{4}$$

When nuclei are under the influence of B_0, the nuclear magnetic moment will precess around the field axis with a certain frequency, υ_0, that is called Larmor frequency, related to B_0 by Equation 5.

$$\nu_0 = \frac{\gamma B_0}{2\pi} \tag{5}$$

Considering ΔE and υ_0, a relation between these parameters can be established by Equation 6, which describes the resonance condition: when the radiation frequency applied by the spectrometer exactly matches ΔE, a transition occurs and a signal will be observed in the NMR

spectrum. Typically, the radiation frequency needed for proton transition is in the radiofrequency region.

$$\Delta E = \frac{h\nu_0}{2} \quad (6)$$

Even though molecules are under a strong external magnetic field, their nuclei are surrounded by moving electrons from chemical bonds that generate a local magnetic field, B_{local}. Consequently, nuclei will have a specific Larmor frequency, which slightly differs from the Larmor frequency expected only when considering the B_0 effect. υ_0, the applied magnetic field frequency of B_0, is in the MHz frequency region whereas the specific Larmor frequency of each nucleus, υ_i, differs only in a dozen or a hundred of Hz. This slight difference is what makes NMR so powerful because nuclei of the same nature in different chemical environments will possess different υ_i and therefore will not be equivalent. This difference is given by the chemical shift (δ) in parts per million (ppm) that is described by Equation 7. By convention, an NMR plot is represented with δ increasing from right to left.

$$\delta = \frac{\nu_i - \nu_0}{\nu_0} \cdot 10^6 \quad (7)$$

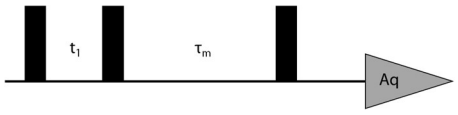
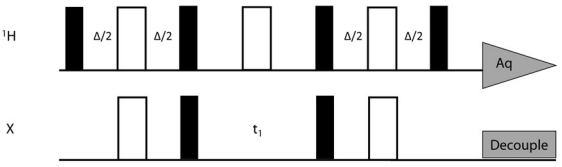
The magnetic field of each nucleus contributes to the local magnetic field experienced by its neighbors, slightly modifying their resonance frequencies. The strength of this interaction is related to the scalar product of the magnetic moments of interacting nuclei by a proportionality constant called the coupling constant – ^nJ (n indicating the number of intervening covalent bonds) – which is measured in Hz and is independent of the applied magnetic field [5].

The most basic NMR experiment corresponds to a one-dimensional (1D) spectrum. The simplest 1D pulse sequence consists of a preparation and a detection period. In the preparation period, a 90º radiofrequency pulse is applied, which rotates the magnetization equilibrium, M, from the z axis (aligned with B_0) to the xy plane. During the detection period, the magnetization will return to the equilibrium (z axis) and its loss along the xy plane will be detected in the form of a free induction decay (FID). The 1D spectrum provides valuable information for small organic molecules and is frequently used for metabolomics [6]. However, for proteins, its usefulness is limited since proteins possess several nuclei which overcrowd the spectra and difficult its interpretation. Therefore, in order to improve the spectral resolution, NMR experiments are acquired with more than one dimension, typically as two dimensions (2D) or three dimensions (3D).

In a 2D experiment, an additional period called the evolution time, containing a variable time delay t_1, is inserted between the preparation and detection periods. The evolution delay increases systematically from zero to the final defined value of increments. After the end of the experiment, all the acquired FIDs are transformed with the same phase parameters. The transformation of the second FID obtained by the t_1 evolution time generates a second frequency domain that results in a 2D spectrum in which the third dimension corresponds to the signal's intensity. These experiments may also contain other periods in addition to the

evolution time such as the mixing period τ_m. 2D NMR experiments can be homonuclear (usually ^{1}H-^{1}H) with a squared-shape spectrum with diagonal peaks, or heteronuclear (usually with a ^{15}N or ^{13}C second frequency dimension) which are asymmetric spectra. The 2D experiments used in this thesis are summarized in Table 2.3.

Table 2.3 - 2D NMR experiments used in this work. Adapted from [7, 8].

NMR Experiment	Information obtained	Pulse sequence[1]
^{1}H,^{1}H-TOCSY	Proton J-couplings within a coupled spin system	
^{1}H,^{1}H-NOESY	Through space (< 5 Å) correlations between dipolarly-coupled spins	
^{1}H,X-HSQC	One-bond heteronuclear coupling between ^{1}H and heteroatom X	

[1] Black and white rectangles represent 90º and 180º radiofrequency pulses, respectively. The gray triangle represents the data acquisition period. The time delay and the mixing time periods are represented by t_1 and τ_m, respectively. DIPSI-2 is a pulse sequence designed for isotropic mixing purposes.

2D ^{1}H,^{1}H-TOCSY (Total Correlation SpectroscopY) experiment correlates all protons within a given spin system. Using J-coupling, magnetization is transferred along the side chain during the mixing time not only for nuclei which are directly coupled but also between nuclei which are connected by a chain of couplings [5].

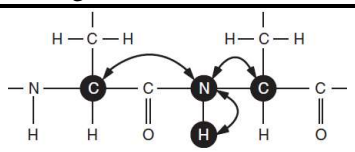
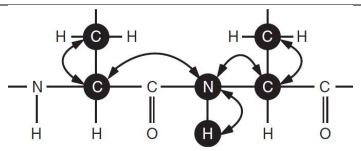
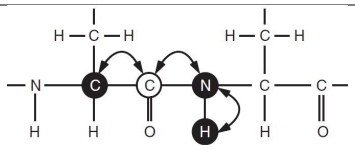
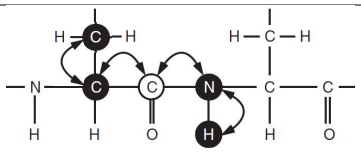
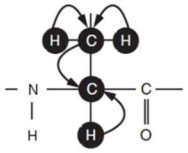
2D ^{1}H,^{1}H-NOESY (Nuclear Overhauser Effect SpectroscopY) experiment relies on the Nuclear Overhauser Effect (NOE), which is defined as the change in intensity of a given nucleus when the energy levels corresponding to a spatially close nucleus are saturated. The NOE arises by cross-relaxation of dipolarly-coupled spins during the τ_m period and the signal intensities in the spectrum are inversely correlated to the sixth power of the distance (r^{-6}) between the interacting spins [5].

In the 2D ^{1}H,X-HSQC, each signal corresponds to single-bond correlations between ^{1}H and a nucleus X (^{15}N or ^{13}C for example). In this experiment, a first pulse sequence generates single quantum coherences between ^{1}H, the most sensitive nucleus, and its attached nucleus and magnetization of both nuclei evolves during t_1. Then, a second pulse sequence is used to transfer the magnetization back to ^{1}H for detection [5, 8].

Three-dimensional (3D) NMR experiments are constructed from 2D NMR experiments by inserting an additional evolution time (t_2) and a second mixing period between the first mixing

period and the acquisition period. Each of the different evolution time periods (t_1, t_2) is incremented separately. Triple resonance experiments possess high sensitivity because magnetization is efficiently transferred via ^1J or ^2J couplings [7]. Table 2.4 describes the 3D NMR experiments used in this thesis.

Table 2.4 - 3D NMR experiments used in this work. Subscripts i and (i-1) indicate a given and its previous amino acid, respectively. Adapted from [8, 9].

NMR experiment	Observed correlations	Magnetization transfer[1]
HNCA	$^1H_i - ^{15}N_i - ^{13}C_{\alpha i}$ $^1H_i - ^{15}N_i - ^{13}C_{\alpha(i-1)}$	
HNCACB	$^{13}C_{\beta i}/^{13}C_{\alpha i} - ^{15}N_i - ^1H_i$ $^{13}C_{\beta(i-1)}/^{13}C_{\alpha(i-1)} - ^{15}N_i - ^1H_i$	
HN(CO)CA	$^1H_i - ^{15}N_i - ^{13}C_{\alpha(i-1)}$	
HN(CO)CACB	$^1H_i - ^{15}N_i - ^{13}C_{\beta(i-1)}/^{13}C_{\alpha(i-1)}$	
HCC(H)-TOCSY	$^1H_{aliphatic} - ^{13}C_{aliphatic}$	

[1]Black circles indicate the nuclei detected by the experiments. White circles indicate nuclei that participate in magnetization transfer but that are not detected in the experiments.

In 3D HNCA, ^1H_i-$^{15}N_i$ nuclei are correlated with $^{13}C_{\alpha i}$ and $^{13}C_{\alpha(i-1)}$. Magnetization originates on amide ^1H nucleus and is transferred to the directly attached ^{15}N. Then, it is transferred via N-C_α J-coupling to $^{13}C_{\alpha i}$ and $^{13}C_{\alpha(i-1)}$ and back again to ^1H for detection. 3D HN(CO)CA is similar to 3D HNCA but only correlates ^1H_i-$^{15}N_i$ nuclei with $^{13}C_{\alpha(i-1)}$, involving a magnetization transfer via $^{13}C_{\alpha(i-1)}$-^{13}CO J-coupling to ^{15}N, although ^{13}CO is not detected in the experiments [8].

3D HNCACB correlates ^1H_i-$^{15}N_i$ nuclei with $^{13}C_{\alpha i}$, $^{13}C_{\beta i}$ and $^{13}C_{\alpha(i-1)}$, $^{13}C_{\beta(i-1)}$. Magnetization is firstly transferred from $^1H_\alpha$ and $^1H_\beta$ to $^{13}C_\alpha$ and $^{13}C_\beta$, respectively, and then secondly from $^{13}C_\beta$ to $^{13}C_\alpha$ for both amino acids. $^{13}C_{\alpha i}$ and $^{13}C_{\alpha(i-1)}$ magnetization is transferred to ^1H_i for detection via $^{15}N_i$. 3D HN(CO)CACB has a similar setup to 3D HNCACB but only correlates ^1H_i-$^{15}N_i$ to $^{13}C_{\alpha(i-1)}$, $^{13}C_{\beta(i-1)}$, including a magnetization transfer step from ^{15}N to $^{13}C_{\alpha(i-1)}$ via ^{13}CO [8].

3D HCC(H)-TOCSY provides connectivities of the aliphatic side chains of individual amino acid residues. In this experiment, magnetization is initiated at ^1H and is then transferred to its

attached ^{13}C and from these to ^{13}C nuclei connected by a chain of couplings. Finally, magnetization is transferred back to ^{1}H for detection [8].

2.4.2 NMR Experiments in the Reduced State

For experiments in the reduced state, samples contained between 0.1 and 1.5 mM of PpcA in sodium phosphate buffer 45 mM with NaCl (100 mM final ionic strength), pH 7.1 either in 92% H_2O/8% $^{2}H_2O$ – for wild-type PpcA solution structure calculation and to observe F6L and W45M backbone amide signals – or in $^{2}H_2O$ (99 % atom) for the mutant's heme substituent signal assignments. For samples prepared in $^{2}H_2O$, the protein was previously lyophilized and suspended in the appropriate buffer prepared exclusively on $^{2}H_2O$.

In order to fully reduce the samples, the NMR tubes were first sealed with a gas-tight serum cap and the air was flushed out to avoid sample reoxidation. Reduction was achieved in the presence of gaseous hydrogen and catalytic amounts of hydrogenase from *Desulfovibrio vulgaris* (Hildenborough), as previously described [10]. The pH of the samples was adjusted prior to its reduction using small amounts of $NaO^{2}H$ or ^{2}HCl and was confirmed after the reduction in an anaerobic LABstar glovebox (MBraun).

2D ^{1}H,^{1}H-TOCSY (20 and 60 ms) and 2D ^{1}H,^{1}H-NOESY (80 ms) were acquired for solution structure calculation. For F6L and W45M mutants, 2D ^{1}H,^{1}H-TOCSY (60 ms) and 2D ^{1}H,^{1}H-NOESY (80 ms) in $^{2}H_2O$ were acquired for heme substituent signal assignments and 2D ^{1}H,^{1}H-TOCSY (60 ms), 2D ^{1}H,^{1}H-NOESY (80 ms) and 2D ^{1}H,^{15}N-HSQC were acquired for backbone amide signal attribution. NMR acquisition parameters for each spectrum are summarized in Table 2.5.

Table 2.5 - NMR acquisition parameters for each experiment in the reduced state. WT indicates wild-type PpcA, MT stands for Mixing Time and NS indicates the number of scans.

	NMR Experiments	MT (ms)	Data points			Spectral width (Hz)			Frequency offset (Hz)			NS
			^{1}H	^{1}H	^{15}N	^{1}H	^{1}H	^{15}N	^{1}H	^{1}H	^{15}N	
WT	**^{1}H,^{1}H-TOCSY**	20	2048	512	-	8417	8417	-	2817	2817	-	64
	^{1}H,^{1}H-TOCSY	60	2048	512	-	8417	8417	-	2817	2817	-	112
	^{1}H,^{1}H-NOESY	80	2048	512	-	8417	8417	-	2817	2817	-	160
Mutants	**^{1}H,^{1}H-NOESY ($^{2}H_2O$)**	80	4096	512	-	8417	8417	-	2819	2819	-	128
	^{1}H,^{1}H-TOCSY ($^{2}H_2O$)	60	2048	512	-	8417	8417	-	2819	2819	-	128
	^{1}H,^{1}H-NOESY	80	2048	512	-	8417	8417	-	2812	2812	-	136
	^{1}H,^{1}H-TOCSY	60	2048	512	-	8417	8417	-	2812	2812	-	136
	^{1}H,^{15}N-HSQC	-	2048	-	128	8417	-	4865	2825	-	8209	160

2.4.3 NMR Experiments in the Oxidized State

All samples for NMR experiments in the oxidized state were prepared containing between 0.4 and 1.5 mM of PpcA in sodium phosphate buffer 45 mM with NaCl (100 mM final ionic strength), pH 5.5 either in 92% H_2O/8% ^2H_2O for assignment of backbone and side chain signals or in ^2H_2O (99 % atom) for assignment of the heme substituent signals. For samples prepared in ^2H_2O, the protein was previously lyophilized and suspended in the appropriate buffer prepared exclusively on ^2H_2O. The pH values of the samples were confirmed with a glass micro electrode.

For backbone and side chain signal assignments, the following spectra were acquired using (1) ^{13}C, ^{15}N-labeled PpcA or (2) natural abundance PpcA: (1) 2D ^1H,^{15}N-HSQC, 2D ^1H,^{13}C-HSQC, 3D HNCACB, 3D HN(CO)CACB, 3D HNCA, 3D HN(CO)CA, 3D HCC(H)-TOCSY and (2) 2D ^1H,^1H-TOCSY (20 and 60 ms) and 2D ^1H,^1H-NOESY (80 ms). For assignment of the heme substituent signals, a 2D ^1H,^{13}C-HSQC, 2D ^1H,^1H-TOCSY (60 ms) and 2D ^1H,^1H-NOESY (80 ms) were acquired using a natural abundance sample. A summary of the NMR acquisition parameters for each experiment are described in Table 2.6.

Table 2.6 - NMR acquisition parameters for each experiment in the oxidized state. MT stands for Mixing Time and NS indicates the number of scans. The + signal indicates data points of 40 (F2) and 128 (F1) in the ^{13}C dimension.

Heteronuclear Experiments	Data points			Spectral width (Hz)			Frequency offset (Hz)			NS
	^1H	^{15}N	^{13}C	^1H	^{15}N	^{13}C	^1H	^{15}N	^{13}C	
HNCACB	2048	40	128	10822	2797	13581	2819	7297	6790	24
HN(CO)CACB	2048	40	128	10822	2797	13581	2819	7297	6790	24
HNCA	2048	40	128	10822	2797	8450	2819	7297	9355	8
HN(CO)CA	2048	40	128	10822	2797	8450	2819	7297	9355	8
HCC(H)-TOCSY	2048	-	+	10822	-	13581	2819	-	6790	16
^1H,^{13}C-HSQC	2048	-	256	28846	-	21127	2819	-	10563	24
^1H,^{13}C-HSQC (^2H_2O)	4096	-	256	27173	-	52824	2823	-	3772	400
^1H,^{15}N-HSQC	2048	256	-	10822	2797	-	2819	7297	-	8

Homonuclear Experiments	MT (ms)	Data points		Spectral width (Hz)		Frequency offset (Hz)		NS
		^1H	^1H	^1H	^1H	^1H	^1H	
^1H,^1H-TOCSY	60	2048	512	27173	27173	2821	2821	208
^1H,^1H-NOESY	80	4096	512	27173	27173	2821	2821	184
^1H,^1H-TOCSY (^2H_2O)	60	2048	512	27173	27173	2821	2821	208
^1H,^1H-NOESY (^2H_2O)	80	4096	512	27173	27173	2802	2802	184

2.4.4 pH Titrations in the Reduced and Oxidized State

In order to probe pH-linked conformational changes of wild-type PpcA in the reduced and oxidized state, the amide signals of the backbone and side chains were followed by 2D $^{1}H,^{15}N$-HSQC experiments in a pH range of 5.5 to 9.6. In the reduced state, the pH of the samples was adjusted with small amounts of $NaO^{2}H$ or ^{2}HCl inside an anerobic glovebox LABstar (MBraun) with O_2 levels below 0.5 ppm to avoid sample oxidation. 2D $^{1}H,^{15}N$-HSQC spectra were acquired collecting 2048 (t_2) x 128 (t_1) data points to cover a spectral width of 7812 (^{1}H) and 4865 (^{15}N) Hz with 8 scans per increment.

For pH titrations in the oxidized state, the sample's pH was adjusted with minimal amounts of NaOH or HCl. 2D $^{1}H,^{15}N$-HSQC spectra were acquired collecting 2048 (t_2) x 128 (t_1) data points to cover a spectral width of 15625 (^{1}H) and 2797 (^{15}N) Hz with 8 scans per increment.

The effects of pH changes in the amide signals were probed using the weighted average chemical shift differences ($\Delta\delta_{average}$) that were calculated as previously described for PpcA from *Gs* (Equation 8):

$$\Delta\delta_{average} = \sqrt{\frac{\Delta\delta^2{}_H + \frac{\Delta\delta^2{}_N}{25}}{2}} \quad (8)$$

In this equation, $\Delta\delta_H$ and $\Delta\delta_N$ correspond to the differences in ^{1}H and ^{15}N chemical shifts, respectively [11].

2.4.5 Assignment Methodologies

The characterization of a protein by NMR encompasses the attribution of its polypeptide signals as a starting point for structural and functional studies such as the solution structure determination or the definition of interacting regions of the protein with relevant partners.

Since PpcA is a cytochrome, NMR characterization not only requires the assignment of the polypeptide signals but also the assignment of the heme substituents' signals. Furthermore, PpcA may exist either in the reduced or oxidized state and, therefore, its characterization must be carried out in both redox states.

2.4.5.1 Backbone and Side Chain Resonance Assignment

Backbone and side chain assignments for PpcA in the reduced state have been previously obtained [3]. In the present thesis the assignment of backbone and side chain signals in the oxidized state was obtained.

The standard NMR backbone assignment methodology typically includes analysis of 3D experiments that correlate ^{1}H-^{15}N signals in the 2D $^{1}H,^{15}N$-HSQC spectrum to the respective

(or sequential) C_α and C_β. In this work, a set of four spectra: 3D HNCACB, 3D HN(CO)CACB, 3D HNCA, 3D HN(CO)CA were analyzed for this purpose.

In 2D ^{1}H,^{15}N-HSQC spectrum, signals will be observed for backbone amides (excluding proline, which does not possess NH due to its cyclic nature, and the N-terminal residue that is in fast exchange with the solvent), side chain amides of asparagine and glutamine and side chain amines of arginine, lysine, histidine and tryptophan residues. Hence, this spectrum constitutes the protein's *fingerprint* because most backbone signals, which define the protein's structure, are represented.

For the backbone assignment, a $^{1}H_i$-$^{15}N_i$ pair is selected from 2D ^{1}H,^{15}N-HSQC spectrum. In the 3D experiments (Figure 2.2), for this $^{1}H_i$-$^{15}N_i$ plane, comparison between 3D HNCA and 3D HN(CO)CA distinguishes $^{13}C_{\alpha i}$ from $^{13}C_{\alpha(i-1)}$. Extending this analysis to 3D HNCACB and 3D HN(CO)CACB permits identification of $^{13}C_{\beta i}$, $^{13}C_{\beta(i-1)}$ and confirms the attributed $^{13}C_{\alpha i}$, $^{13}C_{\alpha(i-1)}$ nuclei. Having found all chemical shifts of a given amino acid i, the next $^{1}H_{i+1}$-$^{15}N_{i+1}$ pair may be found by searching for the ^{1}H-^{15}N plane in which $^{13}C_{\alpha i}$ and $^{13}C_{\beta i}$ are now $^{13}C_{\alpha(i-1)}$ and $^{13}C_{\beta(i-1)}$ in 3D HN(CO)CACB and 3D HN(CO)CA spectra. Other strategy may be to find the previous $^{1}H_{i-1}$-$^{15}N_{i-1}$ pair by searching for the ^{1}H-^{15}N plane in which $^{13}C_{\alpha(i-1)}$ and $^{13}C_{\beta(i-1)}$ are $^{13}C_{\alpha i}$ and $^{13}C_{\beta i}$ in 3D HNCACB and 3D HNCA spectra. Repeating this analysis to every ^{1}H-^{15}N signal in 2D ^{1}H,^{15}N-HSQC permits the complete assignment of all signals.

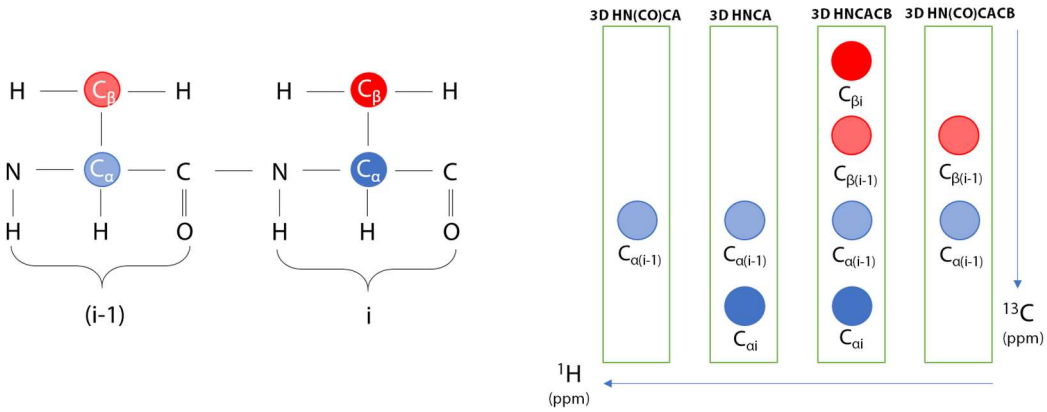


Figure 2.2 – Illustration of 3D NMR spectral features that permit the specific assignment of alpha and beta carbons. The "strips" representing each spectrum are all in the same ^{15}N frequency.

After identification of the ^{1}H, ^{15}N, $^{13}C_\alpha$ and $^{13}C_\beta$ resonances for a given amino acid, the known $^{13}C_\alpha$ and $^{13}C_\beta$ chemical shifts from the backbone assignment are used as a starting point to identify the remaining aliphatic ^{1}H and ^{13}C signals in 3D HCC(H)-TOCSY, which are independently confirmed by analysis of the 2D ^{1}H,^{13}C-HSQC.

2.4.5.2 Heme Substituents Assignment

PpcA has three low-spin heme groups axially coordinated by two histidine residues. Each heme group contributes with 20 proton signals to the NMR spectra – four methyl groups, four meso protons, two thioether methine protons, two thioether methyls and eight protons from two propionate groups (Figure 2.3). Therefore, 60 extra resonances are to be assigned besides those of the 70 residues that comprise the polypeptide chain. Moreover, given the small size of PpcA, the polypeptide chain is mainly confined to the protein's heme core which constitutes an excellent probe to detect overall conformational changes.

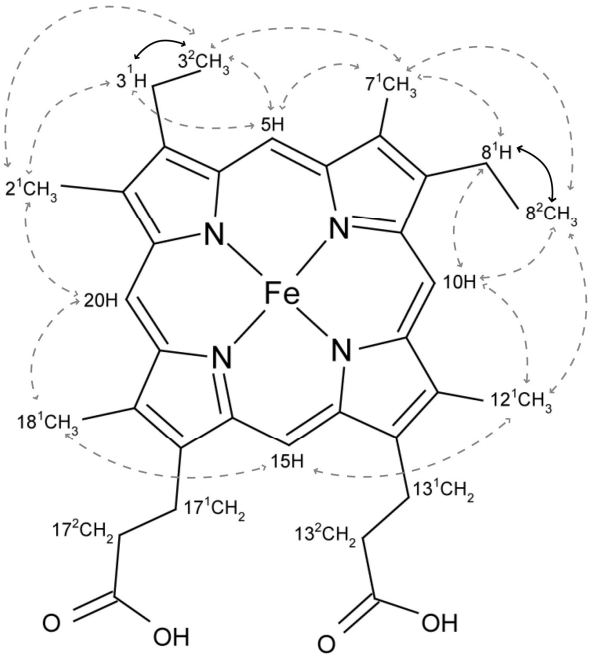


Figure 2.3 - Diagram of a heme *c* numbered according to the IUPAC-IUB nomenclature [12]. Arrows indicate the heme proton connectivities used to assign the heme substituent signals. Solid arrows illustrate connectivities observable in the 2D ^{1}H,^{1}H-TOCSY and that constitute the starting point for the assignment. Dashed arrows correspond to additional NOE connectivities that can be identified in the 2D ^{1}H,^{1}H-NOESY spectrum. In order to not overcrowd the figure, the connectivities involving the heme propionate groups are not represented.

In the reduced state, PpcA is diamagnetic (Fe(II), S=0) and the heme substituents' chemical shifts are essentially dominated by porphyrin ring-current shifts. Consequently, these signals appear in well-defined regions of the NMR spectrum: 10-8 ppm for meso protons (5H, 10H, 15H and 20H); 6-8 ppm for thioether methine protons (3^{1}H and 8^{1}H); 5-2.5 ppm for methyl groups ($2^{1}CH_3$, $7^{1}CH_3$, $12^{1}CH_3$ and $18^{1}CH_3$); and 3 to -1 ppm for thioether methyls ($3^{2}CH_3$ and $8^{2}CH_3$) [13-15]. Heme propionate protons are more variable and therefore there are no particularly well-defined regions attributed to their resonances.

The assignment of the heme substituent signals in the reduced state followed a methodology by Turner and co-workers [16] adapted from the assignment strategy developed by Keller and Wütrich [17] for the horse heart ferricytochrome. First, analysis of the 2D ^{1}H,^{1}H-TOCSY allows

identification of the thioether pairs 3^1H/ 3^2CH$_3$, and 8^1H/8^2CH$_3$ which are the only groups (apart from propionate groups) that are ^{1}J-coupled. In 2D ^{1}H,^{1}H-NOESY spectrum, meso protons present a distinct connectivity pattern according to their location in the heme: 5H and 10H protons are connected to one methyl group and one thioether pair, 15H protons are not connected to any methyl or thioether group and are connected to two propionate groups and 20H protons are connected to two methyl groups (see Figure 2.3). The only ambiguity that arises from these connectivities is between 5H and 10H protons, which is overcome by observing thioether connectivities with 2^1CH$_3$ or 7^1CH$_3$ methyl groups. The specific attribution of each set of heme resonances is made by observing inter-heme connectivities between the heme substituents in the 2D ^{1}H,^{1}H-NOESY spectrum [16].

In the oxidized state, the hemes are paramagnetic (Fe(III), S = 1/2) and the unpaired electron of each heme is responsible for spreading over a large spectral window the heme substituents' signals and those of the amino acids located near the paramagnetic centers. Consequently, the protein's signal attribution is more complex since the typical and distinct spectral regions that are found in the reduced diamagnetic state are no longer observed. This is especially notorious for the heme substituents' signals and, therefore, its assignment is undoubtedly more challenging. However, these obstacles can be potentially overcome by adopting a strategy developed by Morgado and co-workers [18] that combines the analysis of ^{1}H,^{13}C-HSQC NMR spectra obtained for a natural abundance sample and a sample ^{13}C labeled exclusively in its polypeptide chain. This is possible because the heme precursor δ-aminolevulinic acid is not isotopically labeled. Consequently, a short ^{1}H,^{13}C-HSQC (24 scans, considering the protein concentration of 1 mM used in this thesis) acquired for the isotopically labeled protein will not show the heme substituents' signals and, on the contrary, a long ^{1}H,^{13}C-HSQC (400 scans, considering the protein concentration of 0.4 mM used in this thesis) of the natural abundance protein will show them. The superposition of both spectra allows a straightforward identification of most of the heme substituent signals, constituting an important starting point for the assignment. Also, even though the signals show a higher dispersion, there are some characteristic regions: heme methyl substituents (2^1CH$_3$, 7^1CH$_3$, 12^1CH$_3$ and 18^1CH$_3$) typically have higher ^{1}H frequencies and lower ^{13}C frequencies; thioether methines (3^1H, 8^1H) and methyls (3^2CH$_3$, 8^2CH$_3$) typically possess lower and higher ^{13}C frequencies, respectively; propionates α (17^1CH$_2$ and 13^1CH$_2$) and β (17^2CH$_2$ and 13^2CH$_2$) possess lower and higher ^{13}C frequencies, respectively, according to the described by Salgueiro and co-workers [19]. After identification of the hemes' signals, the 2D ^{1}H,^{1}H-TOCSY permitted the assignment of the six pairs of propionates, and the pairs of thioether groups (3^1H/3^2CH$_3$ and 8^1H/8^2CH$_3$). This information, combined with 2D ^{1}H,^{1}H-NOESY, allowed the attribution of each substituent to its specific heme based on the intra- and inter-heme NOE connectivities.

2.4.5.3 Solution Structure Determination in the Reduced State

The solution structure determination is based on the analysis of 2D $^{1}H,^{1}H$-NOESY spectrum in which the signal intensity is inversely proportional to the sixth power of the distance between two interacting ^{1}H spins. Structure determination is an iterative process, driven by the successive identification of new restraints such that the positions of the nuclei are increasingly constrained and lead to a point where the protein's structure is well-defined.

For this, the first step comprehended the assignment of intra and inter-residue NOE connectivities in the 2D $^{1}H,^{1}H$-NOESY [3]. Using the Sparky software, a Gaussian function was used for integration of isolated peaks and a Sum over Box function was used for integration of overlapped peaks. Peaks corresponding to protons separated by fixed distances and all intra-heme cross-peaks were not integrated. The volumes of the NOE cross-peaks were converted into distance restraints by PARADYANA software [20] which calculated a set of 200 conformers and selected the best (lowest energy) 20. These conformers were inspected with Chimera [21] and additional NOE restraints were assigned in the spectrum. For the calculations, non-standard residues were used: fast-flipping aromatic residues with pseudo-atoms to limit the orientations of the planes, proline residues with fixed upper limit distances for ring closure and flexible heme groups [20]. Stereospecific assignments were performed with the help of GLOMSA [22].

2.4.6 Prediction of Secondary Structural Elements with TALOS+

TALOS+ software [23] was used for the prediction of secondary structural elements of PpcA in the oxidized state. The software predicts the protein's ψ and ϕ backbone torsion angles using a combination of ^{1}H, $^{1}H_{\alpha}$, ^{15}N, $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ chemical shifts for a given residue.

2.5 References

[1] Y.Y. Londer, P.R. Pokkuluri, D.M. Tiede, M. Schiffer, Production and preliminary characterization of a recombinant triheme cytochrome c_7 from *Geobacter sulfurreducens* in *Escherichia coli*, Biochimica et Biophysica Acta - Bioenergetics, 1554 (2002) 202-211.
[2] A. Froger, J.E. Hall, Transformation of plasmid DNA into *E. coli* using the Heat Shock Method, Journal of Visualized Experiments, (2007) 252.
[3] P.C. Portela, T.M. Fernandes, J.M. Dantas, M.R. Ferreira, C.A. Salgueiro, Biochemical and functional insights on the triheme cytochrome PpcA from *Geobacter metallireducens*, Archives of Biochemistry and Biophysics, 644 (2018) 8-16.
[4] D.S. Wishart, C.G. Bigam, J. Yao, F. Abildgaard, H.J. Dyson, E. Oldfield, J.L. Markley, B.D. Sykes, ^{1}H, ^{13}C and ^{15}N chemical shift referencing in biomolecular NMR, Journal of Biomolecular NMR, 6 (1995) 135-140.
[5] R.J. Carbajo, J.L. Neira, NMR for Chemists and Biologists, Springer, 2014.
[6] A. Vignoli, V. Ghini, G. Meoni, C. Licari, P.G. Takis, L. Tenori, P. Turano, C. Luchinat, High-Throughput Metabolomics by 1D NMR, Angewandte Chemie International Edition, 58 (2019) 968-994.
[7] Q. Teng, Structural Biology: Practical NMR Applications, Second ed., Springer US, 2013.
[8] J. Cavanagh, W.J. Fairbrother, A.G. Palmer, M. Rance, N.J. Skelton, Protein NMR Spectroscopy: Principles and Practice, Second ed., Elsevier, 2007.
[9] M. Sattlera, J. Schleucher, C. Griesinger, Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution, Progress in Nuclear Magnetic Resonance Spectroscopy, 34 (1999) 93-158.
[10] L. Morgado, M. Bruix, V. Orshonsky, Y.Y. Londer, N.E.C. Duke, X. Yang, P.R. Pokkuluri, M. Schiffer, C.A. Salgueiro, Structural insights into the modulation of the redox properties of two *Geobacter sulfurreducens* homologous triheme cytochromes, Biochimica et Biophysica Acta - Bioenergetics, 1777 (2008) 1157-1165.
[11] D.S. Garrett, Y.-J. Seok, A. Peterkofsky, G.M. Clore, A.M. Gronenborn, Identification by NMR of the Binding Surface for the Histidine-Containing Phosphocarrier Protein HPr on the N-Terminal Domain of Enzyme I of the *Escherichia coli* Phosphotransferase System, Biochemistry, 36 (1997) 4393-4398.
[12] G.P. Moss, Nomenclature of tetrapyrroles, European Journal of Biochemistry, 178 (1988) 277-328.
[13] M. Pessanha, Y.Y. Londer, W.C. Long, J. Erickson, P.R. Pokkuluri, M. Schiffer, C.A. Salgueiro, Redox characterization of *Geobacter sulfurreducens* cytochrome c_7: physiological relevance of the conserved residue F15 probed by site-specific mutagenesis, Biochemistry, 43 (2004) 9909-9917.
[14] M. Pessanha, L. Brennan, A.V. Xavier, P.M. Cuthbertson, G.A. Reid, S.K. Chapman, D.L. Turner, C.A. Salgueiro, NMR structure of the haem core of a novel tetrahaem cytochrome isolated from *Shewanella frigidimarina*: identification of the haem-specific axial ligands and order of oxidation, FEBS Letters, 489 (2001) 8-13.
[15] M.A. Piçarra-Pereira, D.L. Turner, J. LeGall, A.V. Xavier, Structural studies on *Desulfovibrio gigas* cytochrome c_3 by two-dimensional ^{1}H-nuclear-magnetic-resonance spectroscopy Biochemical Journal, 294 (1993) 909-915.
[16] D.L. Turner, C.A. Salgueiro, J. LeGall, A.V. Xavier, Structural studies of *Desulfovibrio vulgaris* ferrocytochrome c_3 by two-dimensional NMR, European Journal of Biochemistry, 210 (1992) 931-936.

[17] R.M. Keller, K. Wüthrich, Assignment of the heme *c* resonances in the 360 MHz ^{1}H NMR spectra of cytochrome *c*, Biochimica et Biophysica Acta (BBA) - Protein Structure, 533 (1978) 195-208.
[18] L. Morgado, A.P. Fernandes, Y.Y. Londer, M. Bruix, C.A. Salgueiro, One simple step in the identification of the cofactors signals, one giant leap for the solution structure determination of multiheme proteins, Biochemical and Biophysical Research Communications, 393 (2010) 466-470.
[19] C.A. Salgueiro, D.L. Turner, A.V. Xavier, Use of paramagnetic NMR probes for structural analysis in cytochrome c_3 from *Desulfovibrio vulgaris*, European Journal of Biochemistry, 244 (1997) 721-734.
[20] L. Brennan, D.L. Turner, A.C. Messias, M.L. Teodoro, J. LeGall, H. Santos, A.V. Xavier, Structural basis for the network of functional cooperativities in cytochrome c_3 from *Desulfovibrio gigas*: solution structures of the oxidised and reduced states, Journal of Molecular Biology, 298 (2000) 61-82.
[21] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, UCSF Chimera—A visualization system for exploratory research and analysis, Journal of Computational Chemistry, 25 (2004) 1605-1612.
[22] P. Güntert, W. Braun, K. Wüthrich, Efficient computation of three-dimensional protein structures in solution from nuclear magnetic resonance data using the program DIANA and the supporting programs CALIBA, HABAS and GLOMSA, Journal of Molecular Biology, 217 (1991) 517-530.
[23] Y. Shen, F. Delaglio, G. Cornilescu, A. Bax, TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts, Journal of Biomolecular NMR, 44 (2009) 213-223.

Chapter 3
Structural and Functional Studies of PpcA in the Reduced State

3 Structural and Functional Studies of PpcA in the Reduced State

The construction of bacterial strains with improved electrogenic capabilities is crucial for the development of sustainable and efficient forms for harvesting energy. In this context, the rational design of proteins with specific reduction potentials is one of the means to achieve this objective. Ueki and co-workers [1] have identified PpcA as crucial for Fe(III) reduction in *Gs*. Furthermore, the measured apparent reduction potential value of an electron transfer reaction in *Gs* attached to electrodes is of -150 mV, a comparable value with the apparent reduction potential of PpcA (-117 mV) [2]. Besides the physiological importance of PpcA, Morgado and co-workers [3] have determined that this protein is neither fully reduced nor fully oxidized in a reduction potential window of -200 mV to -30 mV in a pH range of 5.5 to 9.5, emphasizing its versatility in performing electron transfer. In fact, the dominant microstates in these conditions allow PpcA to couple proton to electron transfer and thus contribute to the proton electrochemical gradient, fundamental for ATP production. Therefore, PpcA constitutes one of the best targets for rational engineering of redox properties. Some studies involving point mutants of this protein have achieved fine-tuning of its redox properties, highlighting the importance of the rational design of proteins with desired redox characteristics [4].

The PpcA family consisting of five triheme *c*-type cytochromes is highly conserved in *Geobacter*. PpcA from *Gs* has been extensively studied and shares 80% of identity with PpcA from *Gm*, differing only in thirteen residues. Studies have suggested that, even though PpcA from *Gm* has a similar structure to PpcA from *Gs* [5], the reduction potentials as well as the order of oxidation of the hemes are markedly different [6]. Thus, this constitutes an excellent system for studying how small changes impact the redox properties of a protein, offering insights on the rational design of proteins with finely-tuned redox properties.

In this thesis we have obtained a preliminary solution structure of PpcA from *Gm* and have probed pH-linked conformational changes that will be used as a starting point for the identification of key residues which modulate the redox properties of this protein.

3.1 Results and Discussion

3.1.1 Solution Structure Determination

The assignment of backbone, side chain and heme substituent signals of PpcA from *Gm* was previously obtained by our group [5] and is available in the Biological Magnetic Resonance Bank (BMRB) entry 27363. Thus, the solution structure determination started with the identification of individual spin-systems and then the assignment of inter-residue NOE cross-peaks in the 2D ^{1}H,^{1}H-NOESY spectrum, according to the described in section 2.4.5.3. The observed HN, H_α and H_β proton connectivities are represented in Figure 3.1.

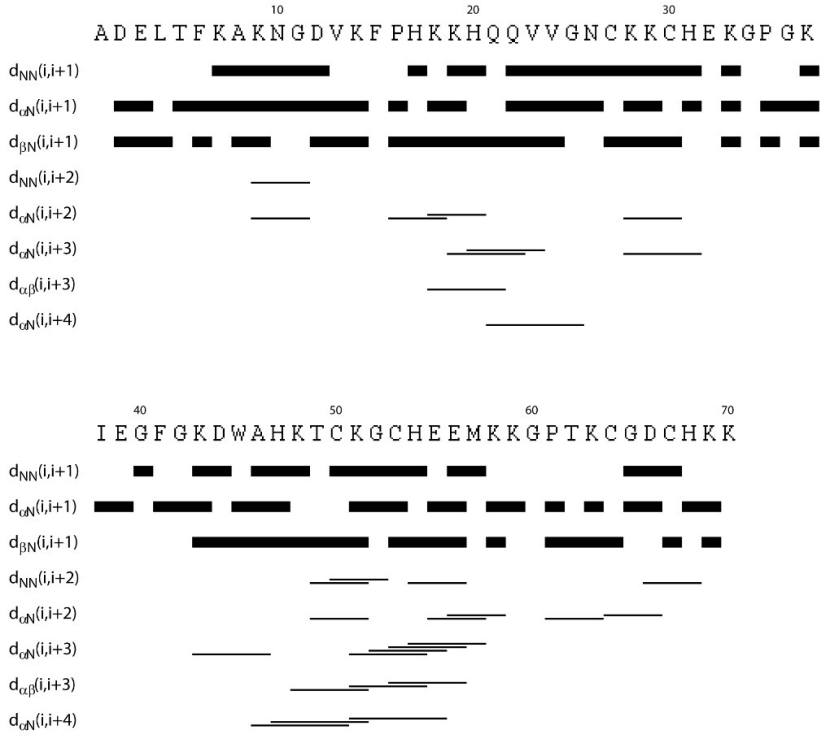


Figure 3.1 - Sequential NOE connectivities involving HN, H_α and H_β protons observed in the 2D ^{1}H,^{1}H-NOESY spectrum of PpcA in the reduced state. The thickness of the bars indicates the intensity of the NOE signals.

The assigned cross-peaks were integrated and converted into volume restraints resulting in a total of 2251 restraints – 830 lov (lower limits for volumes) and 1421 upv (upper limits for volumes) (Table 3.1). These values were used as an input for the program PARADYANA together with a set of 69 fixed upper limit distances (associated with ring closure in the flexible proline residues and heme groups, and the attachment of His axial ligands) [7]. The preliminary structures were analyzed using the program GLOMSA modified to take NOE volumes as input [8] and 10 stereospecific assignments were made for diastereotopic pairs of protons or methyl groups.

Table 3.1 - Summary of restraint violations and quality analysis of the current family of solution structures of PpcA in the reduced state.

Parameter	
Type of distance restraint	
Intra-residue	743
Sequential	521
Medium-Range ($2 \leq \|i - j\| < 5$)	362
Long-Range ($\|i - j\| \geq 5$)	625
Total	2251
	(830 lov + 1421 upv)
Stereospecific Assignments[a]	10
Precision	
Average pairwise RMSD[b] backbone (Å)	0.34 ± 0.08
Average pairwise RMSD heavy atoms (Å)	0.94 ± 0.10

[a]Analysis with GLOMSA [8]

[b]Root Mean Square Deviation

Currently, there is an average of 31 NOE restraints per amino acid (11 lov and 20 upv) and 150 NOE restraints per heme residue (55 lov and 95 upv). The distribution of the number of restraints per residue (Figure 3.2) is not even along the protein as the heme groups in positions 30, 53 and 67 show more long-distance restraints than the remaining amino acids. Although in a lesser extent, residues Leu[4], Phe[6], Val[13], Phe[15], Ile[38] and Lys[43] also show many long-range restraints. It is worth noting that these residues are all located in the proximity of heme groups and that, with exception of Lys[43], are all part of the hydrophobic core, which explains the higher number of long-distance restraints compared to amino acids more exposed to the solvent.

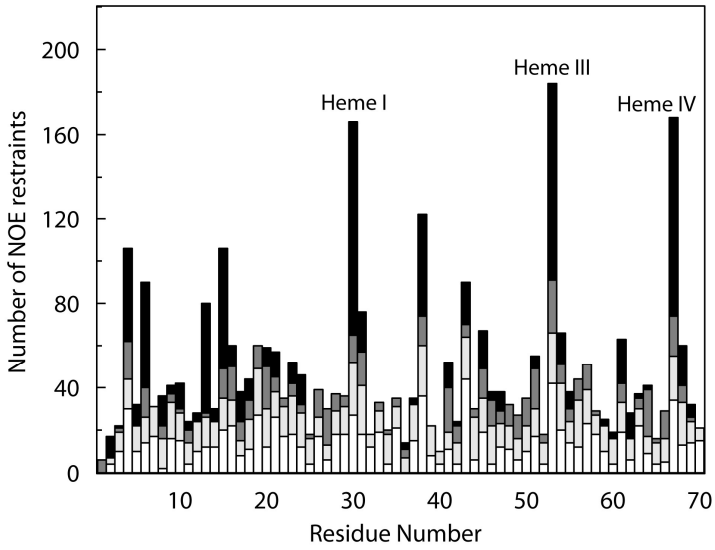


Figure 3.2 - Number of NOE restraints per residue used for the calculation of the current structure of PpcA from *Gm*. White, light gray, dark gray and black bars indicate, respectively, intra-residue, sequential, medium and long-range restraints. Residues 30, 53 and 67 include restraints to hemes I, III and IV.

3.1.2 Analysis of PpcA Structure

The solution structure of PpcA from *Gm* is currently under refinement and the structures of the 20 conformers with lowest target function values are represented in Figure 3.3 with a mean average backbone (NH-C_α-CO) RMSD of 0.34 Å and a heavy atom RMSD of 0.94 Å (Table 3.1). A higher RMSD (Figure 3.4) is observed in the structurally less defined N- and C-terminal regions as well as in the amino acid regions Lys^{28}-Phe^{41} and Glu^{55}-Pro^{65} which form a random coil, and therefore are also associated with higher structural variability.

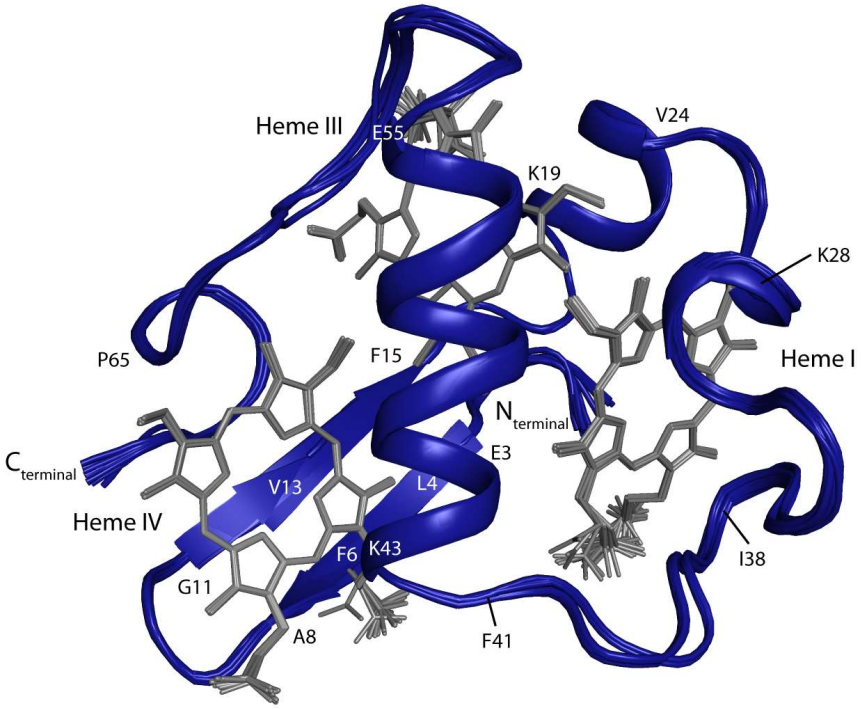


Figure 3.3 - Overlay of the 20 lowest-energy NMR solution structures of PpcA in the reduced state. The polypeptide chain is shown as a blue C_α ribbon and the heme groups are colored gray. Figure produced with PyMOL [9].

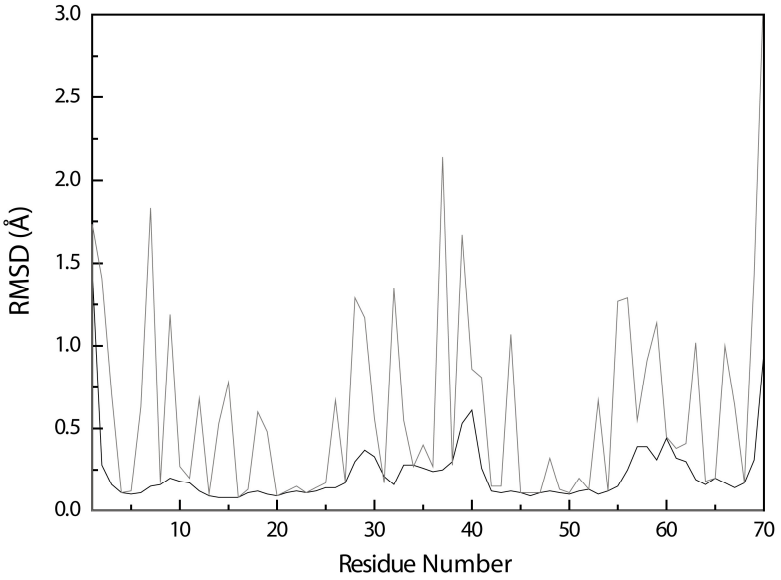


Figure 3.4 - Average pairwise backbone (black) and heavy atom (gray) RMSD values per residue of the family of 20 conformers for the solution structure of PpcA.

The current PpcA structure possesses a β-sheet between residues Glu[3]-Ala[8] and Gly[11]-Phe[15] and two α-helixes between residues Lys[19]-Val[24] and Lys[43]-Glu[55]. Overall, these motifs match those predicted by TALOS+ [10] (Figure 3.5) with exception of the amino acid segment Lys[43] to Met[57] for which TALOS+ predicted a random coil between Gly[52] and His[54] and an α-helix between Glu[55] and Met[57].

Figure 3.5 - Comparison of secondary structural motifs of PpcA given by TALOS+ [10] and from structure calculations of this work. Arrows indicate β-sheet secondary structural motifs whereas helixes represent α-helix secondary structures. Numbers denote the residues involved in the motifs. Adapted from [5].

3.1.3 Pinpointing the Differences Between PpcA from *Gm* and PpcA from *Gs*

The main goal of the determination of PpcA's structure is the elucidation of the structural elements that may explain the functional differences between this protein and PpcA from *Gs*. A preliminary characterization of PpcA from *Gm* [5] showed that the overall pattern of the NH signals in the 2D $^{1}H,^{15}N$-HSQC NMR spectra was similar to that of PpcA from *Gs*, suggesting major structural differences in the Ala[46]-Glu[56] amino acid segment, followed by minor structural differences between His[20] and Lys[29] residues (Figure 3.6A). In solution, PpcA from *Gs* folds in a two-strand β-sheet at the N-terminal formed by Asp[3]–Leu[6] and Val[13]–Pro[16] strands, followed by

three α-helices between residues Ala[19]-Lys[22], Lys[43]-His[47], and Lys[52]-Met[58] [11]. Superposition of the lowest-energy structure of each cytochrome yields a global RMSD difference for backbone atoms of 1.31 Å. Overall, the proteins' global fold is similar (Figure 3.6B), in accordance with the prediction from the analysis of the NMR signals in the 2D ^{1}H,^{15}N-HSQC. The major structural differences are in the Ala[46]-Glu[56] polypeptide segment, particularly between Lys[48] and Lys[51] for which there is no defined secondary structure in PppA from *Gs*, but that is in α-helix in PppA from *Gm*. In this segment, given that sequences only differ in the amino acid at position 49, the structural differences may be explained by two specific amino acids: the first concerns the existence of an extra glycine residue in PppA from *Gs* (*cf.* the amino acid sequence alignment in Figure 3.6C), in position 48, which provides more conformational flexibility to the α-helix segment [12]; the second is related to the existence, in position 45, of a tryptophan in *Gm* rather than a methionine in *Gs*. The former is a bulkier amino acid which may introduce further constraints in the structure. Smaller rearrangements in the α-helix from Lys[19]-Lys[29] are also verified which are in accordance with the prediction solely based on 2D ^{1}H,^{15}N-HSQC comparison (Figure 3.6A).

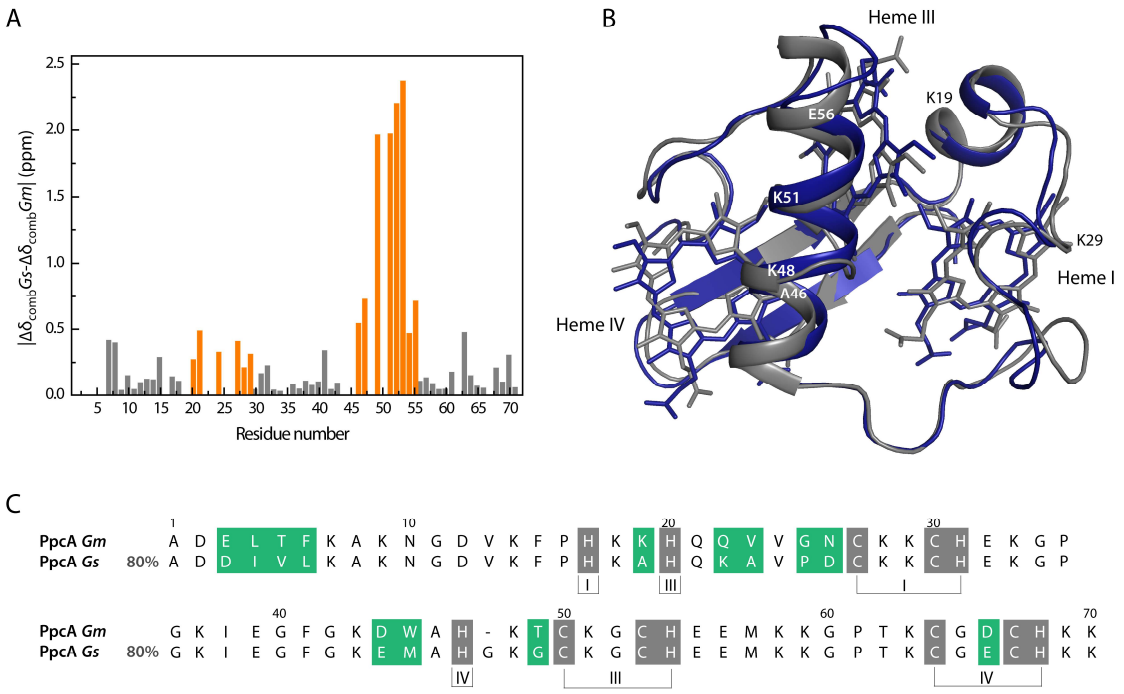


Figure 3.6 – Structural differences between PpcA cytochromes from *Gm* and *Gs*. (A) Differences between the combined ^{1}H and ^{15}N chemical shifts observed for the conserved residues of *Gs* ($\Delta\delta_{comb}Gs$) and *Gm* ($\Delta\delta_{comb}Gm$). The differences were calculated using the equation $\Delta\delta_{comb}=[(\Delta\delta H)^2+ w_i (\Delta\delta N)^2]^{1/2}$, where $\Delta\delta H$ is the difference between ^{1}H shifts, $\Delta\delta N$ the difference between ^{15}N shifts, and $w_i = |\gamma\ ^{15}N\ |/|\gamma\ ^{1}H|$ is a weighting factor that accounts for the differences in nuclei sensitivity [13]. The orange bars highlight the most affected residues. (B) Superposition of the lowest-energy structures of PpcA from *Gm* (blue) and PpcA from *Gs* (PDB code 2LDO) (gray). Polypeptide chain is shown as C_α ribbon. Structures were superposed in PyMOL [9]. (C) Sequence alignment between PpcA from *Gm* and PpcA from *Gs*. The heme binding motifs are colored gray and the non-conserved residues are colored green.

The hemes' disposition is conserved in both proteins with slight rearrangements as observed by the protein's superposition (Figure 3.6B). Comparison between the heme substituents' chemical shifts in both proteins (Figure 3.7A) shows that meso $5H^{III}$ and thioether methyl $3^2CH_3^{III}$ are those with largest chemical shift differences, followed by thioether 3^1H^I and methyl $12^1CH_3^I$ [5]. In a lesser extent, but still showing some differences, are thioether protons 3^1H^{III} and 8^1H^I, thioether methyl $8^2CH_3^I$, methyl $7^1CH_3^{III}$ and meso protons $5H^I$ and $10H^I$. All these substituents are located near the polypeptide segments that present the most conformational rearrangements and the majority of them point toward the protein's hydrophobic core (Figure 3.7B) in which there are two extra aromatic amino acids – Phe^6 and Trp^{45} – which are bulkier (203 Å^3 and 240 Å^3, respectively [14, 15]) than its counterparts in PppA from *Gs* – Leu^6 and Met^{45} – with a volume of 168 Å^3 and 171 Å^3 respectively [14, 15]. The bulkier amino acids will contribute to local conformational changes in the backbone and to the reorientation of nearby residues in the hydrophobic core which will in turn modify the hemes' orientation. This is particularly relevant in the case of these two aromatic amino acids which side chains point toward the cleft between hemes I and III. Also, the introduction of two extra aromatic amino acids will affect the protein's local magnetic field by the local ring currents that will affect the chemical shifts of nearby nuclei [16]. This is particularly evident for $12^1CH_3^I$ and $3^2CH_3^{III}$ heme substituents which establish NOE connectivities with Phe^6 , Phe^{15} and Trp^{45}.

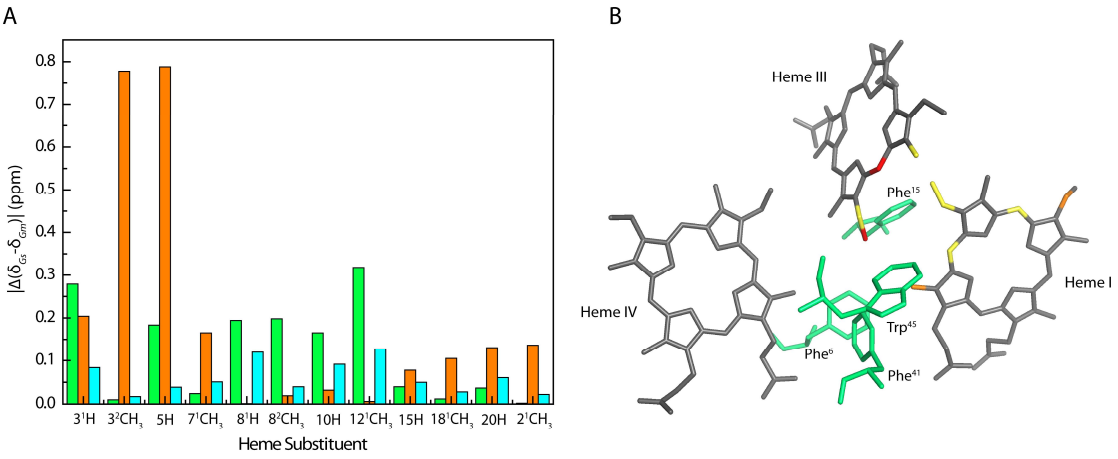


Figure 3.7 – Location of the heme's substituents with higher chemical shift differences between PpcA from *Gm* and PpcA from *Gs*. (A) Chemical shift differences between heme substituents of PpcA from *Gs* (δ_{Gs}) and PpcA from *Gm* (δ_{Gm}). Green, orange and blue bars correspond to hemes I, III and IV, respectively. (B) Mapping of the most affected substituents in the heme core of PpcA from *Gm* lowest-energy structure. The carbon atoms bonded to the most affected protons are colored accordingly to the extent to which the substituents are affected: red [$\Delta(\delta_{Gs} - \delta_{Gm}) > 0.35$ ppm]; orange [0.35 ppm > $\Delta(\delta_{Gs} - \delta_{Gm}) > 0.25$ ppm] and yellow [0.25 ppm > $\Delta(\delta_{Gs} - \delta_{Gm}) > 0.15$ ppm]. Aromatic amino acids are shown in green. Adapted from [5].

The observed structural differences are in agreement with the determined thermodynamic parameters (Table 3.2). In fact, the most prominent differences in heme reduction potentials are in hemes I and III (difference of 74 and 68 mV, respectively, compared to a difference of 12 mV for heme IV) which, as discussed, are the hemes whose substituents showed larger chemical shift

changes and are located in the vicinity of the polypeptide chain with altered conformation. The precise explanation(s) for the differences observed in the reduction potential values of the two cytochromes are still under scrutiny and might be revealed by the solution structure of PpcA from *Gm* coupled to site-directed mutagenesis studies. However, it is possible that the conformational changes observed in the vicinity of hemes I and III may impact their solvent accessibility that is known to influence the reduction potentials [17]. Generally, a decrease in the solvent exposure is associated with more positive reduction potential values and so, in this case, probably heme I and III are less solvent exposed than its counterparts in PpcA from *Gs*.

The thermodynamic characterization of PpcA from *Gm* [6] also showed that its heme reduction potentials are modulated by the solution pH, a property named redox-Bohr effect that allows the protein to associate proton transfer to electron transfer. This modulation is measured by the redox-Bohr interactions (Table 3.2), which are dependent of many factors such as the protonation of heme neighboring residues, electrostatic effects and solvent accessibility [17, 18]. Substitution of Leu^{6} by Phe^{6} in PpcA from *Gs* yielded a mutant with decreased redox-Bohr interactions [19] and, since PpcA from *Gm* possesses a phenylalanine in the same position, it is most likely that this residue contributes to the smaller redox-Bohr interactions. However, only with the definite structure of PpcA can these effects be fully explained.

Table 3.2 - Thermodynamic parameters (in mV) of the fully reduced and protonated forms of PpcA from *Gm* [6] and PpcA from *Gs* [20] at 15 ºC, 250 mM of ionic strength, pH 7. Reduction potentials are relative to NHE.

	Heme reduction potentials			**Heme-heme redox interactions**			**Redox-Bohr interactions**		
	I	III	IV	I-III	I-IV	III-IV	I-H	III-H	IV-H
PpcA *Gs*	-154	-138	-125	27	16	41	-32	-31	-58
PpcA *Gm*	-80	-70	-113	35	3	37	-2	-23	-49

3.1.4 pH-linked Redox Conformational Changes

From the thermodynamic parameters determined for PpcA from *Gm*, Fernandes and co-workers [6] predicted the electron route at pH 7 (Figure 1.7): $p_{0H} \rightarrow p_{4H} \rightarrow p_{14} \rightarrow p_{134}$ where stage *0* is dominated by the protonated microstate p_{0H}, stage *1* is dominated by the microstate p_{4H}, *i.e.* by the oxidation of heme IV while maintaining the redox-Bohr center protonated, stage *2* is dominated by microstate p_{14}, *i.e.* by the oxidation of heme I and deprotonation of the acid-base center, which remains deprotonated in stage *3* in which the oxidation of heme III is dominant (p_{134}). Therefore, at physiological pH, a concerted electron-proton transfer occurs between oxidation stages *1* and *2*.

pH titrations followed by NMR are very useful for monitoring the protonation/deprotonation of the redox-Bohr center since this event will impact the chemical environment of residues in its vicinity and, consequently, their chemical shifts will change accordingly. The pH titration of PpcA from *Gm* was performed in the same experimental conditions as the one for PpcA from *Gs*. Analysis of the chemical shift variation of backbone and side-chain amide signals between pH 5.4 and 9.6 (Figure 3.8) showed that the residues with larger chemical shift variation ($\Delta\delta_{average} > 0.15$ ppm) are backbone amide signals of Glu[3], Lys[7], Ala[8], Lys[18], Ile[38], Trp[45], His[47], Lys[69], and amide side chain signals of Asn[10] and Gln[21]. The affected residues are located in two distinct structural regions: Lys[7], Ala[8], Asn[10], Trp[45], His[47] and Lys[69] are near heme IV whereas Glu[3], Lys[18], Ile[38] and Gln[21] are near heme I.

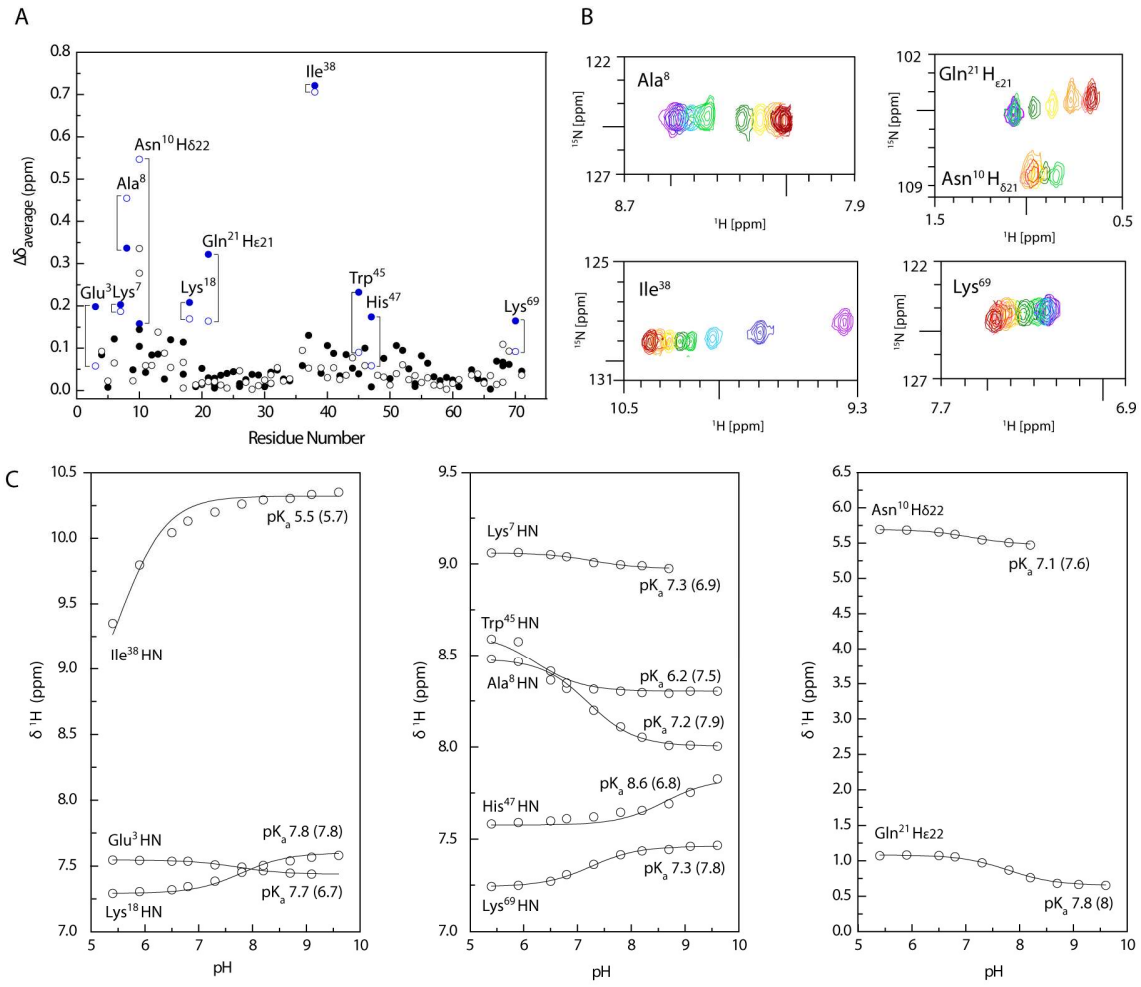


Figure 3.8 - Comparison of pH-linked conformational changes in PpcA from *Gm* (closed circles) and PpcA from *Gs* (open circles). (A) Weighted average of ^{1}H and ^{15}N chemical shifts ($\Delta\delta_{average}$) between pH 5.4 and 9.6. The residues with more significant changes in the $\Delta\delta_{average}$ are colored blue. (B) Expansions of 2D ^{1}H,^{15}N-HSQC for some of the residues with more significant changes in the $\Delta\delta_{average}$. The pH increases from violet to dark red. (C) pH titration curves of the most affected amide signals. The values in parenthesis indicate the pK_a for the correspondent residue in PpcA from *Gs*.

Concerning the amino acids near heme IV (Figure 3.9): Lys[7] (pK_a 7.3), Ala[8] (pK_a 7.2) and Asn[10] (pK_a 7.1) are located in the β-turn segment connecting the two-strand β-sheet near heme IV, Trp[45] (pK_a 6.2) and His[47] (pK_a 8.6) are part of the α-helix that is in between hemes III and IV and Lys[69] (pK_a 7.3) is part of the terminal random coil segment located over heme IV. The NH signals of these residues are facing the propionate 13 of heme IV (P_{13}^{IV}) to which carboxyl group was attributed the redox-Bohr center of this protein [6]. Furthermore, the amino acids' pK_a values are basic (with the exception of Trp[45]) which is in accordance with the determined pK_a of the redox-Bohr center in the fully reduced and protonated form (8.1) by thermodynamic experiments [6], therefore supporting the attribution of this center to P_{13}^{IV}. In the pH titration, the pK_a calculation only considers the variation of the chemical shift of observed signals whereas in the thermodynamic analysis the obtained value reflects the combined contribution of fractional protonation of several acid-base groups [18]. Thus, these differences explain the divergence between the determined pK_a values by the pH titrations and the value obtained from the thermodynamic experiments.

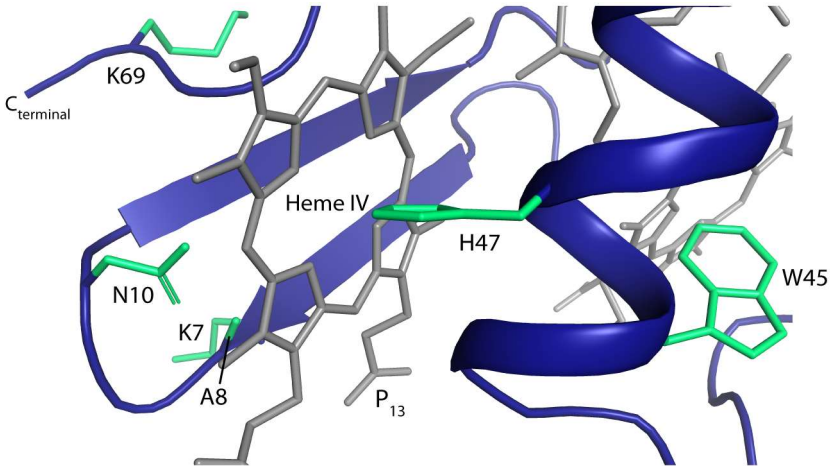


Figure 3.9 – Spatial disposition of residues with significant chemical shift variation on the pH titration located in the vicinity of heme IV. The structure represented is PpcA's lowest-energy structure where is also indicated the redox-Bohr center, P_{13}. Figure generated with PyMOL [9].

Regarding the amino acids near heme I (Figure 3.10), Ile[38] is a fully conserved residue within the family of *Gm* and *Gs* c_7 triheme cytochromes that is known to form a conserved hydrogen bond between its backbone amide proton and the carbonyl oxygen of P_{13}^{I} in the PpcA family from *Gs* [21]. Hence, considering that it is a highly conserved amino acid and given the high structural similarity of PpcA from *Gm* with PpcA from *Gs*, it is expected that it also forms a hydrogen bond in the former. This observation is supported by the acidic pK_a of Ile's backbone amide compared to other residues (pK_a of 5.5 compared to an average pK_a of 7.4 for the remaining residues). Therefore, it is highly likely that the decrease of its proton chemical shift with the pH is due to the disruption of the mentioned hydrogen bond upon protonation of P_{13}^{I}. Regarding Glu[3] (pK_a 7.7), Lys[18] (pK_a 7.8) and Gln[21] (pK_a 7.8), the first is located in the initial segment of the β-sheet whereas Lys[18] and Gln[21] are part of the α-helix in the vicinity of heme I. Since these amino acids are located

far from the redox-Bohr center, their pH dependence may be explained by the protonation of nearby residues of which Lys^{18} is the best candidate, not only because it is positioned between Glu^{3} and Gln^{21} but also because these three residues possess a nearly identical pK_a. Interestingly, the pK_a of Lys^{18} is lower than the expected (~10) which may be explained on a pure electrostatic basis. Adjacent to Lys^{18} is Lys^{19} and, at pH 7, the side chains are positively charged. Thus, the presence of two consecutive positive charges facilitates the deprotonation on one of them, explaining the observed lower pK_a.

In PpcA from *Gs*, the residues which present larger $\Delta\delta_{average}$ are backbone Lys^{7} (0.24 ppm), Ala^{8} (0.31 ppm), Asn^{10} (0.34 ppm), Ile^{38} (0.71 ppm) and the side chain of Asn^{10} (0.55 ppm for $H_{\delta 21}$ and 0.28 ppm for $H_{\delta 22}$). Given that the redox-Bohr center is the same in PpcA from *Gm* and *Gs* and that the pK_a determined by the thermodynamic analysis of PpcA from *Gs* (8.6) [20] only differs by 0.4 from that of PpcA from *Gm*, it is expectable that there are common residues to both proteins involved in pH-linked conformational changes. This is in fact true for backbone amide signals of Lys^{7} and Ala^{8}, and for the $H_{\delta 22}$ side-chain proton of Asn^{10}. Backbone signals of Glu^{3}, Lys^{18}, Trp^{45}, His^{47} and Lys^{69} and the $H_{\varepsilon 22}$ amide signal of Gln^{21} present considerable $\Delta\delta_{average}$ variations that are not present in homologous positions in PpcA from *Gs* (Figure 3.8A).

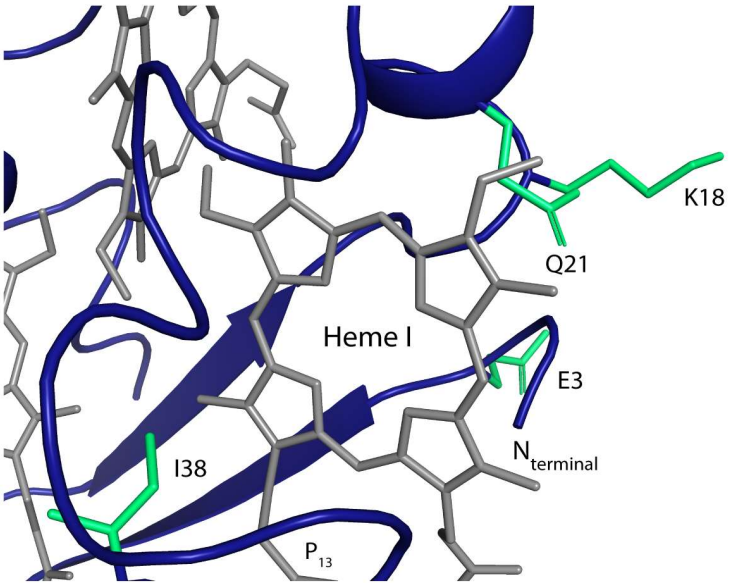


Figure 3.10 - Spatial disposition of residues with significant chemical shift variation on the pH titration located in the vicinity of heme I. The structure represented is PpcA's lowest-energy structure**.** Figure generated with PyMOL [9].

Regarding the common residues to both proteins, in PpcA from *Gs* Lys^{7} establishes a hydrogen bond with the carbonyl oxygen of P_{13}^{IV} whereas Ile^{38} establishes a hydrogen bond with the carbonyl oxygen of P_{13}^{I} [21]. Therefore, given they are conserved amino acids and since the $\Delta\delta_{average}$ is similar in both residues, it is expected that the same intramolecular bonds are established in PpcA from *Gm*, thus explaining the similarities in their pH-dependence. Concerning the conserved residue Ala^{8}, the $\Delta\delta_{average}$ is inferior in PpcA from *Gm* which might be explained by local

rearrangements in the polypeptide chain that locate Ala[8] farther from the redox-Bohr center in PpcA from *Gm*. This hypothesis may also apply to backbone and side chain signals of Asn[10].

Concerning the amino acids which present considerable $\Delta\delta_{average}$ variations only in PpcA from *Gm*, these are located in the amino acid segments with higher structural differences (*cf.* section 3.1.3) which may impact these residues' chemical shifts due to the different chemical environment. Furthermore, the amino acid segment Lys[19]-Lys[29] contains many non-conserved amino acids in both proteins (*cf.* Figure 3.6C) which also contribute for the modification of the chemical shift environment.

3.2 Conclusions

The backbone, side chain and heme substituent NMR signals in the reduced state were used as stepping stones to initiate the determination of PpcA's solution structure and to probe pH-linked conformational changes that will permit the rationalization of the functional data obtained from its thermodynamic characterization. The preliminary solution structure of PpcA from *Gm* shows an overall similar fold to that of PpcA from *Gs* with differences regarding secondary structural motifs in the amino acid segments of Lys^{19}-Lys^{29} and Ala^{46}-Glu^{55}, and in the heme's orientation, particularly in heme I and heme III. The existence of two extra aromatic amino acids in positions 6 and 45 most likely contributes to structural rearrangements in the hydrophobic core which affect the structure. The heme solvent exposure may decrease because of these structural rearrangements, explaining the observed more positive heme reduction potential values. pH-linked conformational changes have shown that the most affected amino acids are located in the vicinity of propionate P_{13} of heme IV, confirming its previous attribution as the protein's redox-Bohr center. These results indicate that the basis for functional differences between PpcA from *Gm* and *Gs* relies on subtle structural rearrangements, highlighting the importance of the study of non-conserved amino acids in both proteins to rationalize the observed results.

3.3 References

[1] T. Ueki, L.N. DiDonato, D.R. Lovley, Toward establishing minimum requirements for extracellular electron transfer in *Geobacter sulfurreducens*, FEMS Microbiology Letters, 364 (2017).
[2] S. Srikanth, E. Marsili, M.C. Flickinger, D.R. Bond, Electrochemical characterization of *Geobacter sulfurreducens* cells immobilized on graphite paper electrodes, Biotechnology and Bioengineering, 99 (2008) 1065-1073.
[3] L. Morgado, J.M. Dantas, M. Bruix, Y.Y. Londer, C.A. Salgueiro, Fine tuning of redox networks on multiheme cytochromes from *Geobacter sulfurreducens* drives physiological electron/proton energy transduction, Bioinorganic Chemistry and Applications, 2012 (2012) 9.
[4] J.M. Dantas, L. Morgado, M. Aklujkar, M. Bruix, Y. Londer, M. Schiffer, P.R. Pokkuluri, C.A. Salgueiro, Rational engineering of *Geobacter sulfurreducens* electron transfer components: a foundation for building improved *Geobacter*-based bioelectrochemical technologies, Frontiers in Microbiology, 6 (2015).
[5] P.C. Portela, T.M. Fernandes, J.M. Dantas, M.R. Ferreira, C.A. Salgueiro, Biochemical and functional insights on the triheme cytochrome PpcA from *Geobacter metallireducens*, Archives of Biochemistry and Biophysics, 644 (2018) 8-16.
[6] T.M. Fernandes, L. Morgado, C.A. Salgueiro, Thermodynamic and functional characterization of the periplasmic triheme cytochrome PpcA from *Geobacter metallireducens*, Biochemical Journal, 475 (2018) 2861-2875.
[7] L. Brennan, D.L. Turner, A.C. Messias, M.L. Teodoro, J. LeGall, H. Santos, A.V. Xavier, Structural basis for the network of functional cooperativities in cytochrome c_3 from *Desulfovibrio gigas*: solution structures of the oxidised and reduced states, Journal of Molecular Biology, 298 (2000) 61-82.
[8] P. Güntert, W. Braun, K. Wüthrich, Efficient computation of three-dimensional protein structures in solution from nuclear magnetic resonance data using the program DIANA and the supporting programs CALIBA, HABAS and GLOMSA, Journal of Molecular Biology, 217 (1991) 517-530.
[9] Schrodinger, LLC, The PyMOL Molecular Graphics System, Version 1.8, in, 2015.
[10] Y. Shen, F. Delaglio, G. Cornilescu, A. Bax, TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts, Journal of Biomolecular NMR, 44 (2009) 213-223.
[11] L. Morgado, V.B. Paixão, M. Schiffer, P.R. Pokkuluri, M. Bruix, C.A. Salgueiro, Revealing the structural origin of the redox-Bohr effect: the first solution structure of a cytochrome from *Geobacter sulfurreducens*, Biochemical Journal, 441 (2012) 179-187.
[12] B.X. Yan, Y.Q. Sun, Glycine residues provide flexibility for enzyme active sites, Journal of Biological Chemistry, 272 (1997) 3190-3194.
[13] F.H. Schumann, H. Riepl, T. Maurer, W. Gronwald, K.-P. Neidig, H.R. Kalbitzer, Combined chemical shift changes and amino acid specific chemical shift mapping of protein–protein interactions, Journal of Biomolecular NMR, 39 (2007) 275-289.
[14] F.M. Richards, Areas, volumes, packing, and protein structure, Annual Review of Biophysics and Bioengineering, 6 (1977) 151-176.
[15] G. Baumann, C. Froömmel, C. Sander, Polarity as a criterion in protein design, Protein Engineering, Design and Selection, 2 (1989) 329-334.
[16] J.A.N.F. Gomes, R.B. Mallion, Aromaticity and Ring Currents, Chemical Reviews, 101 (2001) 1349-1384.

[17] A. Dolla, L. Blanchard, F. Guerlesquin, M. Bruschi, The protein moiety modulates the redox potential in cytochromes *c*, Biochimie, 76 (1994) 471-479.
[18] B.M. Fonseca, I.H. Saraiva, C.M. Paquete, C.M. Soares, I. Pacheco, C.A. Salgueiro, R.O. Louro, The tetraheme cytochrome from *Shewanella oneidensis* MR-1 shows thermodynamic bias for functional specificity of the hemes, JBIC Journal of Biological Inorganic Chemistry, 14 (2009) 375-385.
[19] J.M. Dantas, T. Simões, L. Morgado, C. Caciones, A.P. Fernandes, M.A. Silva, M. Bruix, P.R. Pokkuluri, C.A. Salgueiro, Unveiling the structural basis that regulates the energy transduction properties within a family of triheme cytochromes from *Geobacter sulfurreducens*, The Journal of Physical Chemistry B, 120 (2016) 10221-10233.
[20] L. Morgado, M. Bruix, M. Pessanha, Y.Y. Londer, C.A. Salgueiro, Thermodynamic characterization of a triheme cytochrome family from *Geobacter sulfurreducens* reveals mechanistic and functional diversity, Biophysical Journal, 99 (2010) 293-301.
[21] P.R. Pokkuluri, Y.Y. Londer, X. Yang, N.E.C. Duke, J. Erickson, V. Orshonsky, G. Johnson, M. Schiffer, Structural characterization of a family of cytochromes c_7 involved in Fe (III) respiration by *Geobacter sulfurreducens*, Biochimica et Biophysica Acta (BBA)-Bioenergetics, 1797 (2010) 222-232.

Chapter 4
Structural and Functional Studies of PpcA in the Oxidized State[1]

[1] The work presented in this Chapter was partially reproduced from P.C. Portela, J.M. Dantas, C.A. Salgueiro, Backbone, side chain and heme resonance assignment of the triheme cytochrome PpcA from *Geobacter metallireducens* in the oxidized state. Biomolecular NMR Assignments (2019).

4 Structural and Functional Studies of PpcA in the Oxidized State

The pH dependence of reduction potential values (redox-Bohr effect) in PpcA implicates that it possesses the adequate functional properties to perform electron/proton energy transduction. Hence, it might contribute as an additional mechanism for the creation of the proton electrochemical gradient across the periplasmic membrane, essential for ATP synthesis [1, 2]. Understanding the functional mechanism by which PpcA is able to couple proton to electron transfer requires its characterization both in the reduced and oxidized states in order to identify the structural changes and residues that enable it. Moreover, the high identity and different functional properties observed between PpcA from *Gs* and PpcA from *Gm* makes it imperative to thoroughly characterize the latter to understand the observed differences [1, 3].

The structural and functional characterization of PpcA in the reduced state was presented in Chapter 3, and in this Chapter the study was extended to the oxidized state, by assigning the backbone, side chain and heme resonances and by probing pH-linked conformational changes. This initial characterization lays foundations for future studies of PpcA in the oxidized state – such as the solution structure determination and the mapping of interactions with relevant physiological partners – that, altogether, will contribute to the elucidation of extracellular electron transfer processes.

4.1 Results and Discussion

4.1.1 Heme, Backbone and Side Chain NMR Resonance Assignments

In the paramagnetic oxidized state, the observed chemical shifts result from the combination of the heme's ring-current effect, which is also present in the reduced state, and the paramagnetic contribution of the unpaired electrons. These two contributions affect the signals' broadness and are responsible for its spreading over a large spectral window. The signal broadening and the lack of well-defined regions for the heme substituents and nearby residues makes the assignment in the oxidized state much more challenging [4]. The assignment of backbone and side chain signals is facilitated by the existence of well-established strategies that take advantage of the $^{13}C/^{15}N$ enrichment of the sample, which permits the dispersion of signals over more than two dimensions [5]. However, usually the heme precursor δ-aminolevulinic acid is not isotopically labeled, and therefore there is no experiment that permits a straightforward identification of the heme substituents' signals. This obstacle can be potentially overcome by adopting a strategy developed by Morgado and co-workers [6] that combines the analysis of 2D $^{1}H,^{13}C$-HSQC NMR spectra obtained for a natural abundance sample and a sample ^{13}C labeled exclusively in its polypeptide chain as described in section 2.4.5.2. The superposition of both spectra (Figure 4.1) allows a straightforward identification of most of the heme substituent signals, constituting an important starting point for the assignment.

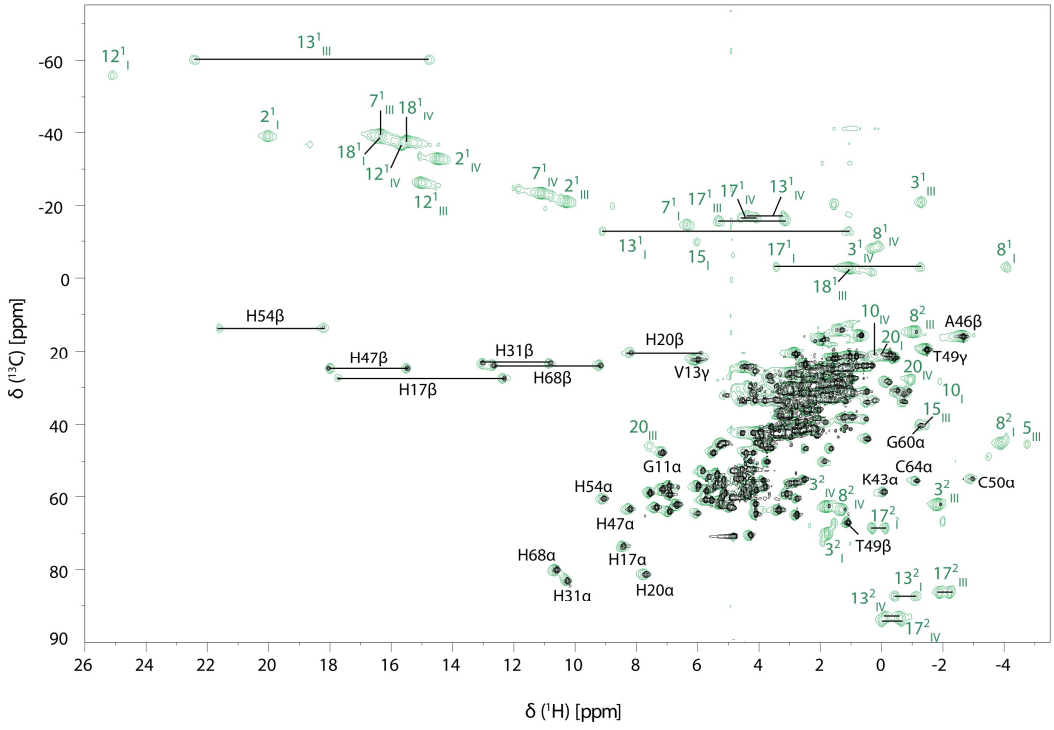


Figure 4.1 - Overlay of 2D $^{1}H,^{13}C$-HSQC NMR spectra of PpcA in the oxidized state. The spectra correspond to PpcA sample labeled ($^{13}C/^{15}N$) exclusively in its polypeptide chain (black contours) and a natural abundance one (green contours). Green and black labels indicate the heme substituent and polypeptide signals, respectively. Protons associated to the same carbon atom are connected by a straight line. Heme substituents are indicated in accordance to the IUPAC-IUB nomenclature for tetrapyrroles [7].

The inter-heme connectivities found in this assignment are represented in Figure 4.2 and are common to those existent in the reduced state [3], suggesting conservation of the hemes' orientation, in accordance with a previous study on the hemes' magnetic properties and axial ligands orientation [8]. Fewer inter-heme NOE connectivities were identified in the oxidized state compared to the reduced state (Figure A1.5 in Appendix A) since some substituents could not be assigned with confidence and others were overlapped. Altogether, it was possible to assign 53 out of 60 of the heme substituents which chemical shifts are reported in Table A2.4 (Appendix A).

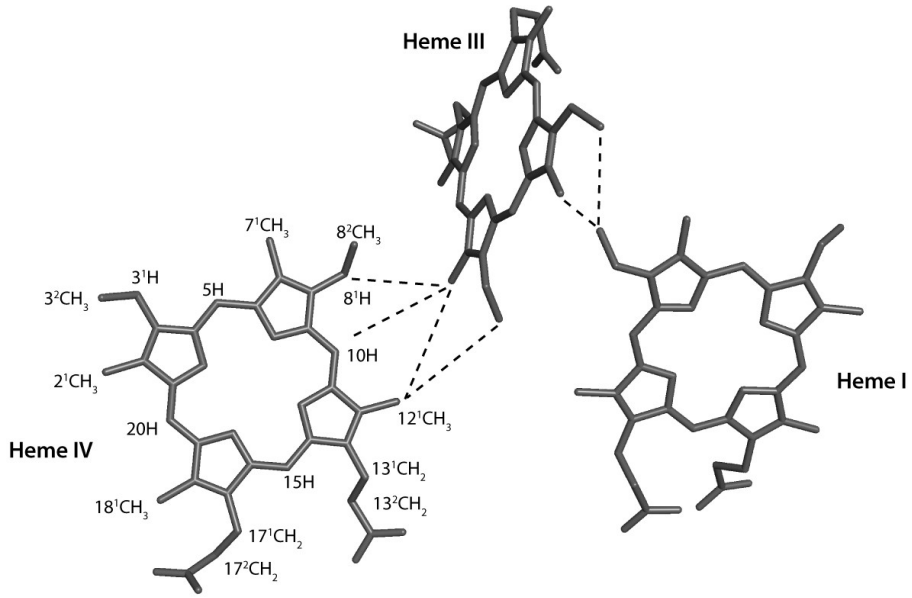


Figure 4.2 - Inter-heme NOE connectivities observed for PpcA in the oxidized state. The represented heme core corresponds to that of the lowest-energy solution structure of PpcA in the reduced state. Black dashed lines indicate the inter-heme NOE connectivities observed in the 2D $^{1}H,^{1}H$-NOESY spectrum of PpcA. The IUPAC-IUB nomenclature for tetrapyrroles [7] is represented on heme IV.

The assignment of the backbone resonances was achieved by integrating the information from 2D $^{1}H,^{15}N$-HSQC, 3D HNCACB, 3D HN(CO)CACB, 3D HNCA, 3D HN(CO)CA which allowed near complete assignment of the signals: ^{15}N (94%), ^{1}HN (94%), $^{13}C_{\alpha}$ (97%) and $^{13}C_{\beta}$ (97%). The analysis of the 3D HCC(H)-TOCSY and the 2D $^{1}H,^{13}C$-HSQC permitted to extend the assignment to the remaining aliphatic protons ($^{1}H_{\alpha}$, $^{1}H_{\beta}$, $^{1}H_{\gamma}$, $^{1}H_{\delta}$ and $^{1}H_{\varepsilon}$) and aliphatic carbons ($^{13}C_{\gamma}$, $^{13}C_{\delta}$ and $^{13}C_{\varepsilon}$). It was possible to assign 87 % of ^{1}H signals (including 99% of $^{1}H_{\alpha}$ and 97% of $^{1}H_{\beta}$), 97% of ^{13}C signals and 92% of ^{15}N signals, excluding carboxyl and hydroxyl groups. The ^{1}H, ^{13}C and ^{15}N chemical shifts of the assignment have been deposited in the BMRB under the accession number 27987 and are reported in Table A2.5 – Appendix A.

The 2D $^{1}H,^{15}N$-HSQC NMR spectrum of a labeled PpcA sample with the backbone nitrogen resonances assigned, as well as those of the side chains of glutamine (Gln^{22}), asparagine (Asn^{26}), and tryptophan (Trp^{45}) residues is shown in Figure 4.3. It was not possible to assign the ^{15}N-^{1}H pairs of Ala^{1}, Asp^{2}, Gly^{36} and Gly^{42} most likely because the amide protons of N-terminal residues

Ala[1] and Asp[2] are in fast-exchange with the solvent [9] and because of broadening of Gly[36] and Gly[42] signals due to their proximity with hemes I and IV, respectively, a situation also verified in the spectrum of PpcA from *Gs*.

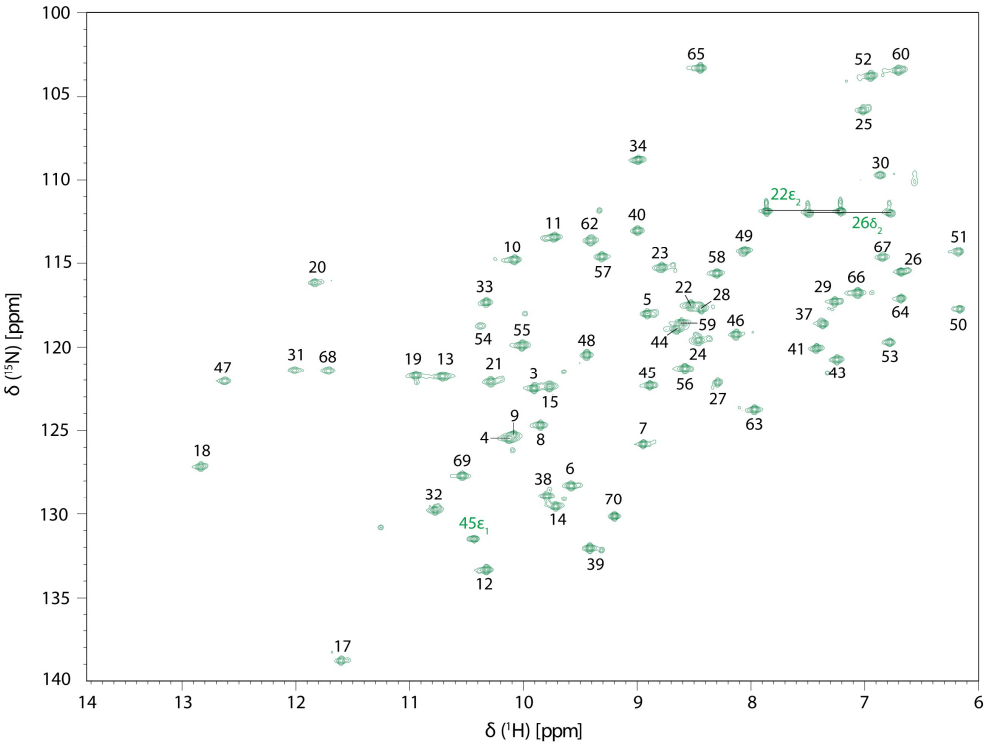


Figure 4.3 - 2D $^{1}H,^{15}N$-HSQC spectrum of PpcA in the oxidized state. Backbone ^{15}N-^{1}H signals are labeled black while side chain resonances (Gln[22] $NH_{\varepsilon 2}$, Asn[26] $NH_{\delta 2}$ and Trp[45] $NH_{\varepsilon 1}$) are labeled green. The protons associated to the same nitrogen atom are connected by a straight line.

Compared to the 2D $^{1}H,^{15}N$-HSQC NMR spectrum of PpcA in the reduced state [3], side chain resonances of Asn[10], Gln[21] and histidine (His[17], His[20], His[31], His[47], His[54], His[68]) residues are not found. In the case of Asn[10] and Gln[21], a possible explanation may be the fact that both residues are located in the vicinity of heme groups (heme IV and I respectively) and most likely become significantly broader and undetectable (*cf.* structural positions of these two residues in Figure 3.9 and 3.10, respectively). Similarly, all histidine residues are heme axial ligands and, as consequence of their close contact with the paramagnetic center, the NH_{ε} signals of their side chain are also considerably broadened [10]. Moreover, Phe[15] aromatic ring signals are not observed, suggesting broadening due to paramagnetic effects, similar to what is observed in the same residue of PpcA from *Gs*. On the other hand, for the aromatic rings of Phe[6] and Phe[41] residues, three well-resolved resonances were identified and indicate that they are in a fast flipping molecular motion regime [11]. It was also possible to assign the seven aromatic resonances of Trp[45].

As referred, besides signal broadening, the paramagnetic effect also causes the spread of resonances of amino acids in the vicinity of the hemes. The most affected signals are those of αCH and βCH_2 of the six axial histidines (His[17], His[20], His[31], His[47], His[54], His[68]), αCH of cysteine

residues that are part of the heme binding motif (Cys^{50} and Cys^{64}), βCH_3 of Ala^{46}, βCH of Thr^{49}, γCH_3 of Val^{13} and Thr^{49}, and αCH of Gly^{11}, Lys^{43}, Gly^{60} (Figure 4.1). Even though some of these signals are located near the ones belonging to the heme substituents, the assignment is facilitated by the direct comparison between the $^{1}H,^{13}C$-HSQC acquired for labeled and natural abundance samples (Figure 4.1).

In order to inspect whether the PpcA structure in the oxidized state resembled that of the reduced state, PpcA secondary structural motifs were predicted by the TALOS+ software [12] and the results are depicted in Figure 4.4. The software predicts two β-sheets between residues Glu^{3}-Lys^{7} and Asp^{12}-Val^{13} and α-helix motifs between residues His^{17}-Val^{24}, Gly^{42}-Lys^{48} and Gly^{52}-Met^{57}. It also predicts isolated secondary structural motifs in amino acids Glu^{32}, Ile^{38} and Cys^{64} which were not considered since a single isolated amino acid cannot be part of a secondary structural motif itself. Overall, these motifs are similar to those of the PpcA reduced structure (*cf.* Figure 3.5) with differences in the β-sheet segment Gly^{11}, Lys^{14}-Phe^{15} and in the α-helix segments His^{17}-Lys^{18}, Thr^{49}-Lys^{51} and Glu^{56}-Met^{57}.

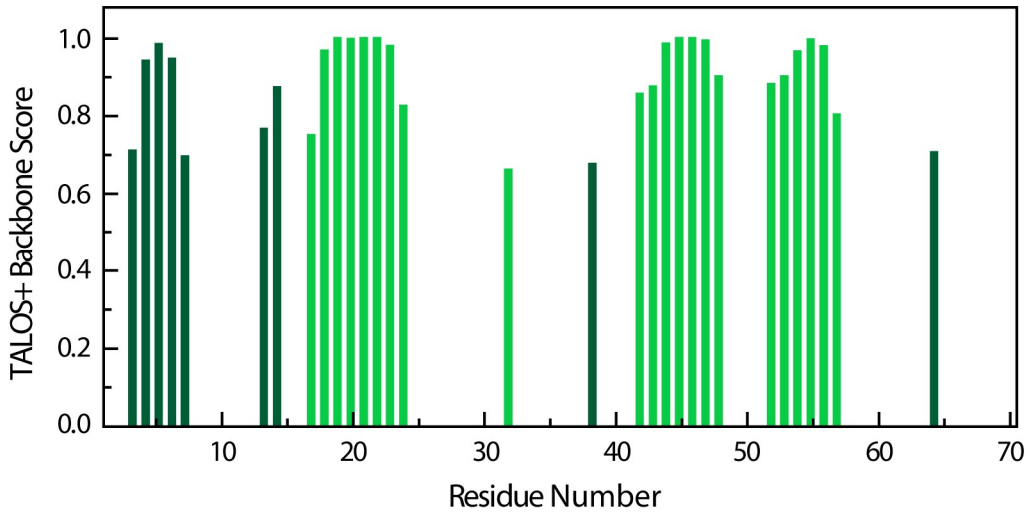


Figure 4.4 - TALOS+ secondary structural motifs prediction for PpcA in the oxidized state. Dark green bars indicate β-sheet structural elements whereas light green bars indicate α-helix structural elements. The TALOS+ Backbone Score indicates the confidence in the predicted secondary structure elements.

The observed differences between the predicted secondary structural elements of the oxidized and reduced states may reflect, effectively, a conformational change in the polypeptide chain and/or the prediction may be biased by some contribution from the paramagnetic shifts. Given the high sequence similarity between PpcA from *Gm* and PpcA from *Gs*, and that the solution structure of the latter has been determined in the oxidized state [13], the information obtained from the comparison between ^{15}N-^{1}H pairs of the 2D $^{1}H,^{15}N$-HSQC spectrum of both proteins can complement the information given by the TALOS+ software and, therefore, assess the likelihood of the prediction.

This comparison benefits from the information that the heme core orientation is conserved and that the axial ligands are nearly in the same orientation as the ones in PpcA from *Gs* [8], meaning that the observed differences will indicate with more confidence the polypeptide segments with different conformation. Figure 4.5A shows that the largest chemical shift differences are in the Ala^{46}-Glu^{55} segment, suggesting an overall similar folding of the polypeptide chain (*cf.* secondary structures in Figure 4.5B) with important structural differences in the mentioned segment. In fact, this polypeptide segment is also the most affected one in the reduced state [3] (corroborated by the preliminary solution structure of PpcA *cf.* Figure 3.6A). Therefore, these observations suggest that the Ala^{46}-Glu^{55} amino acid segment is, indeed, a region with a distinct conformation when comparing secondary structures of different oxidation stages in PpcA, and when comparing the structures of PpcA from *Gm* and PpcA from *Gs*. It is important to notice that some amino acids in the initial β-sheet (Lys^{7}-Phe^{15}), the first α-helix segment (His^{20}-Cys^{27}) and the Lys^{59}-Lys^{63} segment also present, to some extent, differences in the combined chemical shift values suggesting that there might be small conformational differences between both proteins.

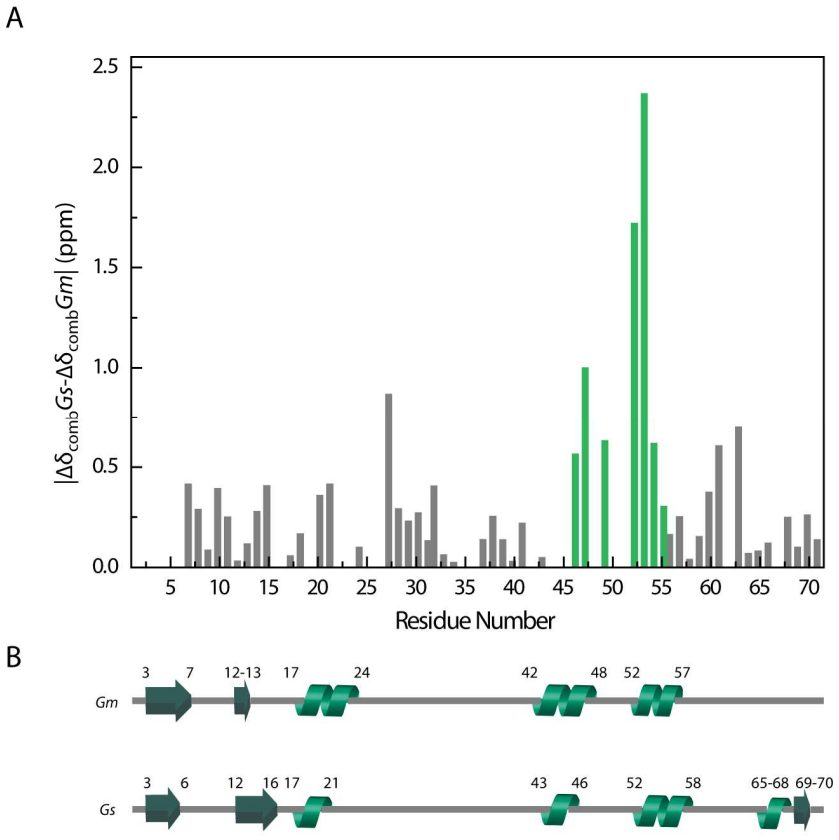


Figure 4.5 - Mapping of structural differences between PpcA from *Gm* and *Gs* in the oxidized state. (A) Differences between the combined ^{1}H and ^{15}N chemical shifts for conserved residues in PpcA from *Gs* ($\Delta\delta_{comb}Gs$) and *Gm* ($\Delta\delta_{comb}Gm$). The differences were calculated using the equation $\Delta\delta_{comb}=[(\Delta\delta H)^2+ w_i (\Delta\delta N)^2]^{1/2}$, where $\Delta\delta H$ is the difference between ^{1}H shifts, $\Delta\delta N$ the difference between ^{15}N shifts, and $w_i = |\gamma\ ^{15}N\ |/|\gamma\ ^{1}H|$ is a weighting factor that accounts for the differences in nuclei sensitivity [14]. The green bars highlight the most affected residues. (B) Secondary structural elements of PpcA from *Gm*, predicted by TALOS+ in this work, and of PpcA from *Gs* (PDB code 2MZ9) [13]. Arrows and helixes indicate β-sheet and α-helix secondary structural elements, respectively.

4.1.2 Probing of pH-linked Redox Conformational Changes

The detailed thermodynamic characterization of PpcA indicated the existence of a redox-Bohr center with different pK_a values for the reduced (8.1) and oxidized (6.5) states [1]. As referred, pH titrations followed by 2D $^{1}H,^{15}N$-HSQC NMR spectra can be explored to identify the redox-Bohr center and residues which might be associated with conformational changes. In this titration, it was not possible to follow the signals of Ala^{1}, Asp^{2}, Gly^{36} and Gly^{42} which were already not observable in the 2D $^{1}H,^{15}N$-HSQC spectrum, as described in section 4.1.1.

The NH signals which present the largest variations in the pH titration ($\Delta\delta_{average} > 0.15$ ppm) are Ala^{8} (pK_a of 6.4 and 8.8), Phe^{15} (pK_a of 6.2), Lys^{18} (pK_a of 7.2), Ile^{38} (pK_a of 6.1) and Cys^{50} (pK_a of 9.1) (Figure 4.6A-C). Regarding the structural location of the residues (Figure 4.6D), Ala^{8} is part of the β-sheet turn and is the residue closest to heme IV propionate 13 (P_{13}^{IV}). Cys^{50} is located in the second α-helix in-between hemes III and IV, with its amide group also pointing toward P_{13}^{IV}. Phe^{15} is part of the β-sheet segment and is in-between hemes I and III whereas Lys^{18} is part of the first α-helix near heme I. Finally, Ile^{38} amide backbone is pointed toward propionate 13 of heme I (P_{13}^{I}). Remarkably, Ala^{8} presents two pK_a values in contrast to the other residues which present only one, suggesting different origins for the pH-linked conformational changes [15]. In PpcA from *Gs*, the residues with largest variations in their $\Delta\delta_{average}$ were Lys^{7}, Ala^{8}, Phe^{15}, Lys^{18}, Ile^{38} and Ala^{46} (see Figure 4.6A). The majority of the affected residues are common to PpcA from *Gm*, with the only difference being in Lys^{7} and Ala^{46} residues, suggesting that the redox-Bohr center is also conserved in these two proteins.

Concerning the amino acids of PpcA from *Gm* that present significant chemical shift variation with the pH, Ala^{8} has a more acidic pK_a value in the oxidized state (6.4) compared to the reduced form (7.2), which matches that of the redox-Bohr center calculated by thermodynamic studies ($pK_a^{ox} = 6.5$) and, given its proximity to P_{13}^{IV}, reiterates the attribution of this group as the redox-Bohr center. The existence of a second pK_a (8.8) may be related to the protonation of nearby residues such as Lys^{7} or Lys^{9} that significantly affect Ala^{8} chemical shift [15]. Even though the observed pK_a is much lower than the typical value for the side-chain of a lysine (~10), the existence of two spatially close, positively charged residues may facilitate the deprotonation of one of them and thus explain the observed pK_a. In addition, in PpcA from *Gs*, the side chain position of Lys^{43} alters its conformation from the reduced to the oxidized state to interact more closely with P_{13}^{IV} [13]. This may also occur in PpcA from *Gm* and may impact the hydrogen bond (H-bond) network in this area, affecting Ala^{8} chemicals shifts. The smaller $\Delta\delta_{average}$ verified for this residue in PpcA from *Gm* may be explained by structural rearrangements in the β-sheet, further confirmed by the comparison between $\Delta\delta_{comb}Gs$ and $\Delta\delta_{comb}Gm$ (Figure 4.5). These structural rearrangements may be, on one hand, the reason why Lys^{7} presents significant chemical shift differences in PpcA from *Gs* but not in PpcA from *Gm* and, on the other, may affect Phe^{15} such that it is more susceptible to the protonation/deprotonation of the redox-Bohr center, thus explaining its more acidic pK_a in comparison with PpcA from *Gs*.

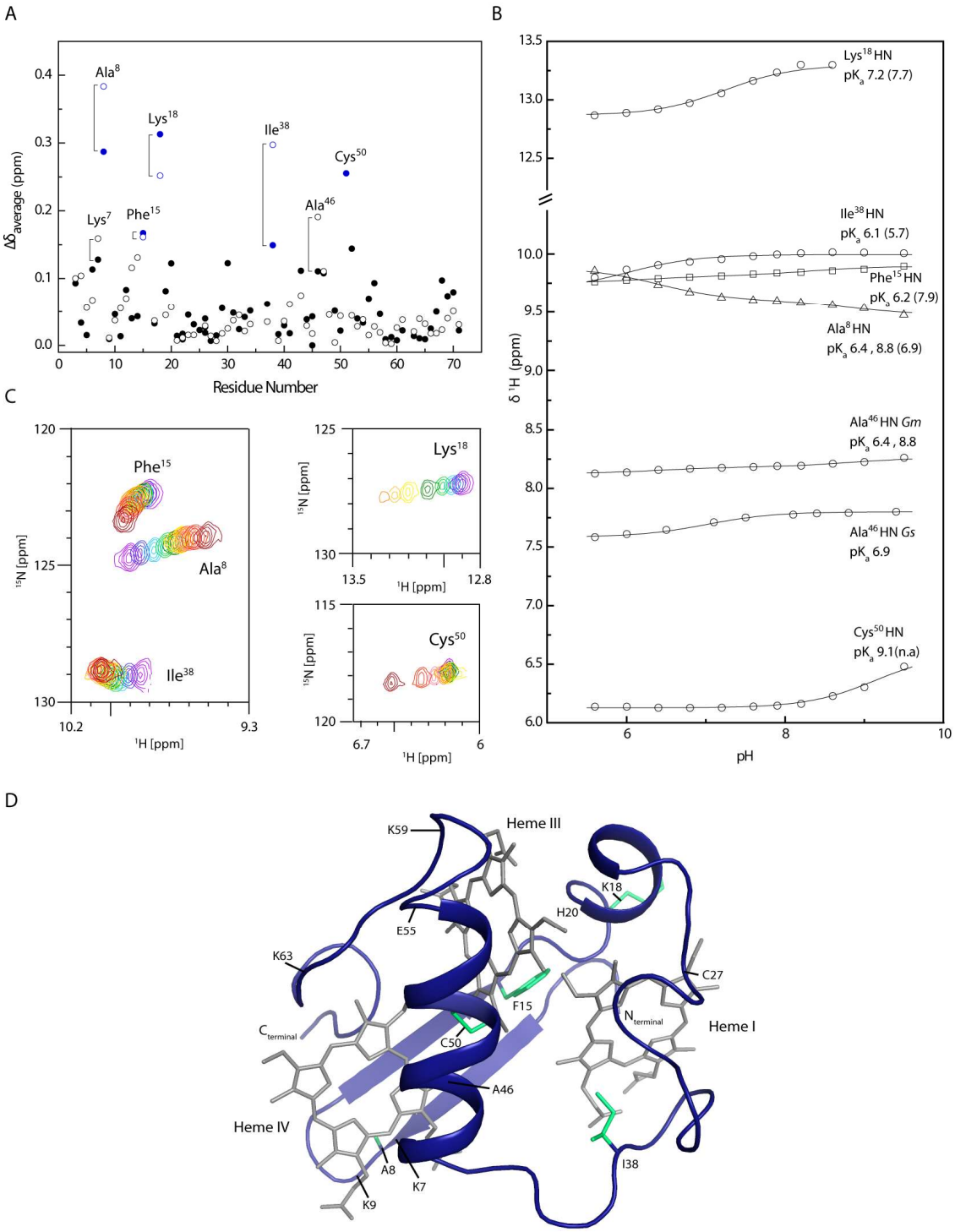


Figure 4.6 - Comparison of pH-linked conformational changes in PpcA from *Gm* (closed circles) and PpcA from *Gs* (open circles). (A) Weighted average of ^{1}H and ^{15}N chemical shifts ($\Delta\delta_{average}$) between pH 5.5 and 9.6. The residues with more significant changes in the $\Delta\delta_{average}$ are colored blue. Lys^{7} and Ala^{46}, which present significant changes in PpcA from *Gs*, are also indicated. Data for PpcA from *Gs* were obtained from Morgado and co-workers [13]. (B) pH titration curves of the most affected amide signals in this titration and of Ala^{46}. The values in parenthesis indicate the pK_a for the correspondent residue in PpcA from *Gs*. (C) Expansions of the 2D ^{1}H,^{15}N-HSQC for the residues with the most significant changes in the $\Delta\delta_{average}$. The pH increases from violet to dark red. (D) Structural location of the affected residues (side chains colored green) in PpcA's lowest-energy structure of the reduced state. Figure generated with PyMOL [16].

In the oxidized structure of PpcA from *Gs*, Phe^{15} does not establish an H-bond with the main chain oxygen of Ile^{4}, being susceptible to the pH variations. Given the conservation of this residue in both proteins, this may occur as well in PpcA from *Gm*, elucidating why Phe^{15} does not present a significant chemical shift variation in the pH titration in the reduced state but is significantly affected in the pH titration of the oxidized state [13, 15].

The basic pK_a value of Cys^{50} is in accordance with the basic pK_a value of Ala^{8}, further supporting the occurrence of an event in the vicinity of hemes III and IV. The residue Ala^{46} is significantly affected in PpcA from *Gs* but not in PpcA from *Gm*. The different conformation of this segment (as suggested by a previous analysis - Figure 4.5) may explain the lower $\Delta\delta_{average}$ in PpcA of *Gm*. The pH dependence of the chemical shifts of Ala^{46} of PpcA from *Gm* shows two pK_a values (6.4 and 8.8) instead of only one value observed in PpcA from *Gs* (Figure 4.6C) further confirming the observations for Ala^{8}.

In PpcA from *Gs*, the backbone NH signal from residue Lys^{18} presented a significant variation only in the oxidized state since in the reduced one it established an H-bond with the main chain oxygen of Pro^{16}. This observation, combined with the fact that Lys^{18} is affected in both redox states in PpcA from *Gm* (*cf.* Figures 3.8A and 4.6A), suggests that this segment presents a different H-bond network that makes Lys^{18} in PpcA from *Gm* more susceptible to pH variation [13].

The more acidic pK_a of Ile^{38} (pK_a of 6.1) may have the same genesis as the reduced state, where the pH dependence was attributed to the disruption of a hydrogen bond between the amide group of Ile^{38} and the carboxylate group of P_{13}^{I} upon protonation of the latter. However, in contrast with the reduced state, the variation in the chemical shift is much less significant than its counterpart in PpcA from *Gs*, hinting different redox-dependent conformational changes in the two cytochromes.

Lastly, residues Glu^{3}, Asn^{10}, Gln^{21}, Trp^{45}, His^{47} and Lys^{69} were significantly affected by the pH in the reduced state but not in the oxidized state. As discussed in the previous and present sections, there may be subtle structural rearrangements between the reduced and oxidized state that modify the residues' conformation and respective H-bond network and thus account for the observed differences. The determination of the solution structure of PpcA in the reduced and oxidized states and their structural comparison will permit the rigorous assessment of the conformational differences here suggested.

4.2 Conclusions

The assignment of backbone, side chain and heme NMR resonances of PpcA in the oxidized state constitutes an important step in its structural and functional characterization. Based on this information, it was possible to predict the secondary structural motifs of this cytochrome which suggest that PpcA maintains its overall structure regardless the oxidation state, with slight conformational changes. Comparison of the 2D $^{1}H,^{15}N$-HSQC spectrum with that of PpcA from *Gs* highlights the amino acid segment Ala^{46}-Glu^{55}, near heme III, as the most different between both proteins. The mapping of pH-linked conformational changes has validated P_{13}^{IV} as the redox-Bohr center and identified common and dissimilar affected residues in the oxidized and reduced states. In addition, compared to the data obtained for PpcA from *Gs*, the more affected residues are not strictly conserved in the two cytochromes, suggesting that the mechanism by which both cytochromes couple proton to electron transfer is not identical.

The obtained results lay foundations for: (i) the determination of the solution structure of PpcA in the oxidized state that will elucidate in detail the observations made in this work and (ii) the mapping of interactions with redox partners to clarify the extracellular electron transfer processes.

4.3 References

[1] T.M. Fernandes, L. Morgado, C.A. Salgueiro, Thermodynamic and functional characterization of the periplasmic triheme cytochrome PpcA from *Geobacter metallireducens*, Biochemical Journal, 475 (2018) 2861-2875.
[2] L. Morgado, J.M. Dantas, M. Bruix, Y.Y. Londer, C.A. Salgueiro, Fine tuning of redox networks on multiheme cytochromes from *Geobacter sulfurreducens* drives physiological electron/proton energy transduction, Bioinorganic Chemistry and Applications, 2012 (2012) 9.
[3] P.C. Portela, T.M. Fernandes, J.M. Dantas, M.R. Ferreira, C.A. Salgueiro, Biochemical and functional insights on the triheme cytochrome PpcA from *Geobacter metallireducens*, Archives of Biochemistry and Biophysics, 644 (2018) 8-16.
[4] C.A. Salgueiro, D.L. Turner, A.V. Xavier, Use of paramagnetic NMR probes for structural analysis in cytochrome c_3 from *Desulfovibrio vulgaris*, European Journal of Biochemistry, 244 (1997) 721-734.
[5] J. Cavanagh, W.J. Fairbrother, A.G. Palmer, M. Rance, N.J. Skelton, Protein NMR Spectroscopy: Principles and Practice, Second ed., Elsevier, 2007.
[6] L. Morgado, A.P. Fernandes, Y.Y. Londer, M. Bruix, C.A. Salgueiro, One simple step in the identification of the cofactors signals, one giant leap for the solution structure determination of multiheme proteins, Biochemical and Biophysical Research Communications, 393 (2010) 466-470.
[7] G.P. Moss, Nomenclature of tetrapyrroles, European Journal of Biochemistry, 178 (1988) 277-328.
[8] T.M. Fernandes, L. Morgado, C.A. Salgueiro, D.L. Turner, Determination of the magnetic properties and orientation of the heme axial ligands of PpcA from *Geobacter metallireducens* by paramagnetic NMR, Journal of Inorganic Biochemistry, 198 (2019) 110718.
[9] Y. Bai, J.S. Milne, L. Mayne, S.W. Englander, Primary structure effects on peptide group hydrogen exchange, Proteins: structure, function, and bioinformatics, 17 (1993) 75-86.
[10] G.N. La Mar, Nuclear Magnetic Resonance of Paramagnetic Macromolecules, 2 ed., Springer, Dordrecht, 1995.
[11] K. Wüthrich, G. Wagner, NMR investigations of the dynamics of the aromatic amino acid residues in the basic pancreatic trypsin inhibitor, FEBS Letters, 50 (1975) 265-268.
[12] Y. Shen, F. Delaglio, G. Cornilescu, A. Bax, TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts, Journal of Biomolecular NMR, 44 (2009) 213-223.
[13] L. Morgado, M. Bruix, P.R. Pokkuluri, C.A. Salgueiro, D.L. Turner, Redox- and pH-linked conformational changes in triheme cytochrome PpcA from *Geobacter sulfurreducens*, Biochemical Journal, 474 (2017) 231-246.
[14] F.H. Schumann, H. Riepl, T. Maurer, W. Gronwald, K.-P. Neidig, H.R. Kalbitzer, Combined chemical shift changes and amino acid specific chemical shift mapping of protein–protein interactions, Journal of Biomolecular NMR, 39 (2007) 275-289.
[15] J.E. Nielsen, Analyzing Protein NMR pH-Titration Curves, in: R.A. Wheeler, D.C. Spellmeyer (Eds.) Annual Reports in Computational Chemistry, Elsevier, 2008, pp. 89-106.
[16] Schrodinger, LLC, The PyMOL Molecular Graphics System, Version 1.8, in, 2015.

Chapter 5
Role of Phe^{6}, Val^{13} and Trp^{45} Residues in PpcA[1]

[1] The work presented in this Chapter was partially reproduced from J.M. Dantas, P.C. Portela, A.P. Fernandes, Y.Y. Londer, X. Yang, N.E.C. Duke, M. Schiffer, P.R. Pokkuluri, C.A. Salgueiro, Structural and functional relevance of the conserved residue V13 in the triheme cytochrome PpcA from *Geobacter sulfurreducens*, The Journal of Physical Chemistry B, 123 (2019) 3050-3060.

The NMR data for Val[13] mutants were already available in the host laboratory. My contribution to this research topic was the NMR resonance assignment and the interpretation of the results, which are presented in this Chapter.

5 Role of Phe[6], Val[13] and Trp[45] Residues in PpcA

Site-directed mutagenesis is an excellent methodology to probe the importance of residues within a protein. With this in mind, some studies have been made in PpcA from *Gs* addressing the strictly conserved residue Phe[15] [1], the Met[58] residue that protects heme III from solvent exposure [2], several lysine residues spatially close to the heme groups [3] and the importance of Leu[6] residue in the modulation of the redox-Bohr effect [4]. The two previous chapters have provided the structural characterization of PpcA from *Gm* in the reduced and oxidized states. This lays the foundations for subsequent studies of the key residues that modulate the redox properties in this protein which we will address in this chapter.

Phe[6], Val[13] and Trp[45] residues are located in the hydrophobic core of the protein and establish NOE connectivities with the heme groups. The site-directed mutagenesis of these residues had as primary objective the assessment of their importance in the structural and functional properties of PpcA in the reduced state. Val[13] is a conserved residue among all cytochromes from the family of PpcA from *Gs* and *Gm* (with exception of PpcE in the latter, in which is replaced by a isoleucine) [5]. The site-directed mutagenesis was used to study the Val[13] impact in the heme core by replacing it with non-polar residues (alanine and isoleucine) or with polar residues (serine and threonine). This study was performed using PpcA from *Gs* due to the existence of available data at the start of this thesis. Phe[6] residue is highly conserved in the PpcA family from *Gm* but not in the family of PpcA from *Gs* and Trp[45] is conserved in three members of the PpcA family from *Gm* and is only present in PpcD of *Gs*. So far, the members of the PpcA family from *Gm* have presented more positive reduction potential values than the proteins belonging to the latter family [5, 6]. Hence, we hypothesized that these residues could be involved in the reduction potential shift to more positive values in PpcA from *Gm* since they are in close contact with the heme groups. To test this hypothesis, we have replaced Phe[6] and Trp[45] residues in PpcA from *Gm* by its counterparts in PpcA from *Gs* – Leu[6] and Met[45].

The studies presented in this chapter constitute a step towards the goal of identifying, from the thirteen residues that differ in PpcA from *Gm* and *Gs*, the key participants in the modulation of the reduction potentials of two functionally different proteins.

5.1 Results and Discussion

5.1.1 Structural Role of Val[13]

One of the hallmarks of the cytochrome PpcA is the highly compact arrangement of the three heme groups within the small hydrophobic core of the protein (Figure 5.1) with an average of 24 residues per heme group. Due to this arrangement, any disturbance in the polypeptide chain will influence the hemes and, therefore, they are excellent probes to inspect the impact of mutations in the protein [7]. PpcA's polypeptide chain folding provides a very particular chemical environment to the heme core so that the observed heme substituent NMR signals are essentially determined by their relative orientation and neighboring residues. In fact, because the hemes are low-spin in the reduced state, typical regions for the heme substituent signals are observed: 8 to 10 ppm for meso protons (5H, 10H, 15H and 20H); 6 to 8 ppm for thioether methines (3^1H and 8^1H); 2.5 to 5 ppm for methyl groups (2^1CH_3, 7^1CH_3, 12^1CH_3 and 18^1CH_3); and -1 to 3 ppm for thioether methyls (3^2CH_3 and 8^2CH_3).

With exception of PpcE of *Gm*, the Val[13] residue is conserved in the PpcA family of both *Gm* and *Gs* and is located within *van der Waals* distance of hemes III and IV (Figure 5.1). In order to probe the effects of side chain and polarity in the heme core, valine was replaced by hydrophobic residues alanine (V13A) and isoleucine (V13I), or by polar residues serine (V13S) and threonine (V13T). Based on the structure of PpcA, it was expected that isoleucine could not be accommodated without a significant perturbation in the local structure and, on the contrary, serine and threonine were expected to be accommodated within the spatial constraints.

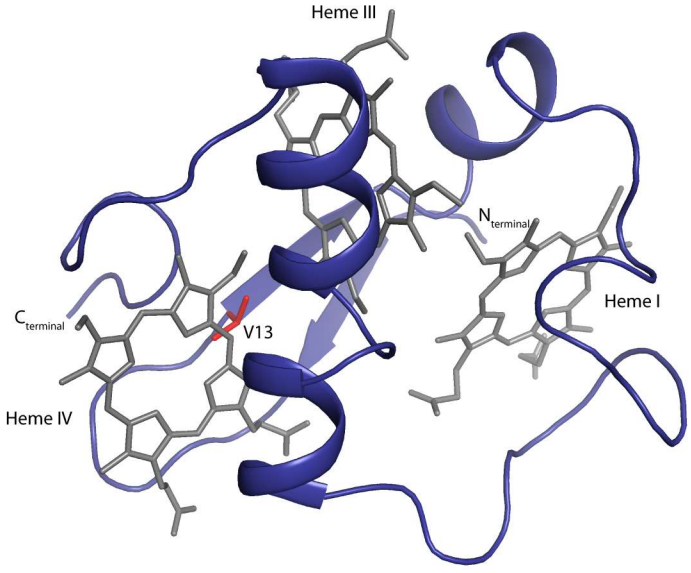


Figure 5.1 - Spatial location of the residue Val[13] in the solution structure of PpcA from *Gs* (PDB code 2LDO [8]). The PpcA polypeptide chain (blue) is shown as C_α ribbon and heme groups are colored gray. The side chain of Val[13] (red) is represented as stick drawings. Figure generated with PyMOL [9].

The 2D ^{1}H NMR spectra (NOESY and TOCSY) were used to assign the heme substituents' signals in each mutant in the reduced state and to inspect the effect of the mutations on the cytochrome's heme core. The signals were assigned using the same methodology described for the wild-type protein [10] which was described in section 2.4.5.2. With the exception of V13A, and in contrast to the wild-type cytochrome, all the mutants showed more than six connectivities in the thioether methine/thioether methyl proton region in the 2D ^{1}H,^{1}H-TOCSY spectra (Figure 5.2). The number of connectivities observed suggests that V13I, V13S and V13T mutants have two conformations for the heme core in solution (hereafter named conformations A and B). The presence of the two conformations was further confirmed in the 2D ^{1}H,^{1}H-NOESY spectra of each mutant and is illustrated by the NOE connectivities between 5H and 10H meso protons and their respective thioether methine protons (Figure 5.2). The assignment of the heme substituents in a multiheme cytochrome showing multiple conformations was achieved for the first time in the present work and the full list of assignments is provided in Table A2.6 of the Appendix A.

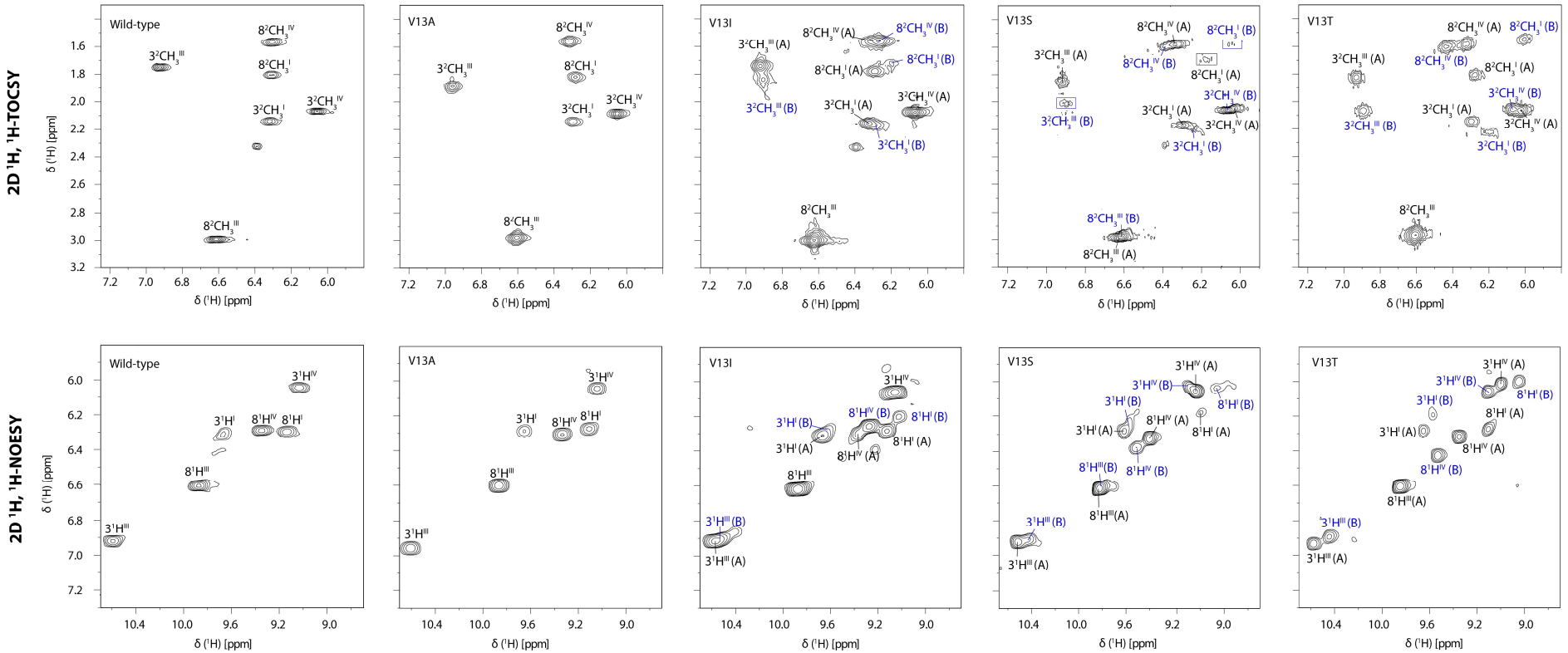


Figure 5.2 - Selected regions of 2D ^{1}H,^{1}H-TOCSY and 2D ^{1}H,^{1}H-NOESY NMR spectra to illustrate the presence of different conformations in solution. The connectivities between the thioether methines (3^{1}H or 8^{1}H) and the thioether methyl groups ($3^{2}CH_3$ and $8^{2}CH_3$) are illustrated in the 2D ^{1}H,^{1}H-TOCSY spectra (upper panels). The NOE connectivities between 5H and 10H meso protons and their respective thioether methine protons (3^{1}H and 8^{1}H) are illustrated in the 2D ^{1}H,^{1}H-NOESY spectra (lower panels). Black and blue labels represent the signals of conformations A and B, respectively.

The comparison between the chemical shifts of the mutated and wild-type cytochromes indicates that the overall heme core is not considerably affected despite displaying minor variations in two conformations (Figure 5.3A). This is supported by the low RMSD values (Figure 5.3A and Table A2.6 in the Appendix A) and by the same set of NOE connectivities observed in the 2D ^{1}H,^{1}H-NOESY NMR spectra for both wild-type and mutants.

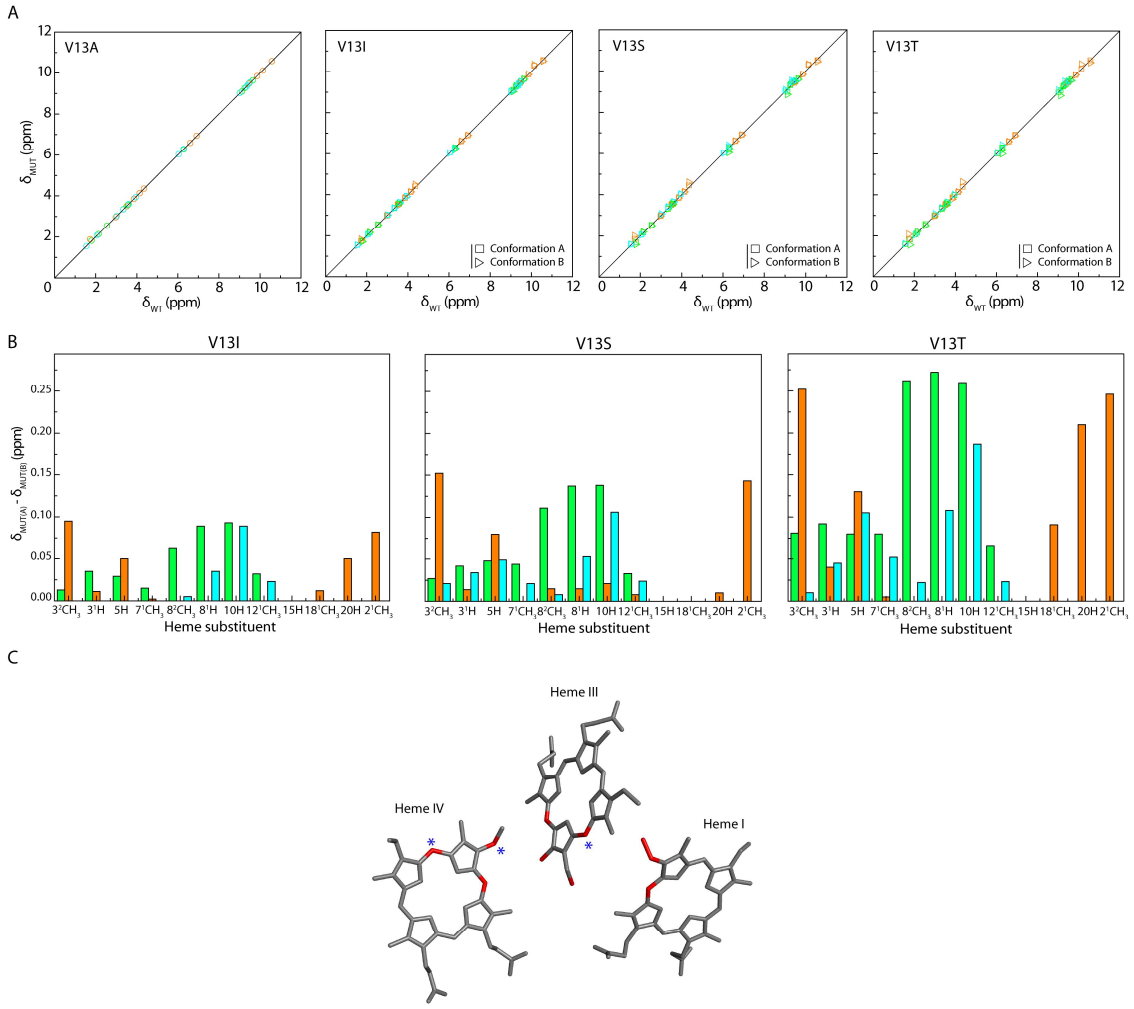


Figure 5.3 - Impact of the mutations V13A, V13I, V13S and V13T on the heme core architecture. (A) Comparison of the heme proton chemical shifts of the PpcA valine mutants (δ_{MUT}) and those of PpcA (δ_{WT}). Green, orange, and blue symbols correspond to hemes I, III, and IV, respectively. The RMSD values calculated from the chemical shifts measured for the wild-type and mutants are: (i) V13A: 0.01, 0.05 and 0.01 ppm for hemes I, III and IV, respectively; (ii) V13I: 0.01(0.05), 0.05(0.08) and 0.02(0.03); (iii) V13S: 0.06(0.13), 0.07(0.13) and 0.05(0.08); (iv) V13T: 0.01(0.15), 0.03(0.15) and 0.03(0.09). The values given in parenthesis correspond to the conformation B. The solid line has a unit slope. (B) Variation of the heme proton chemical shifts in conformations A and B ($\delta_{MUT(A)}$ - $\delta_{MUT(B)}$) for V13I, V13S and V13T. The green, orange, and blue bars correspond to hemes I, III, and IV, respectively. (C) Heme core of PpcA (PDB code 2LDO [8]) highlighting the most affected signals in the two conformations for V13I, V13S and V13T. The carbon atoms bonded to the most affected protons in the three mutants are colored red. Blue asterisks indicate carbon atoms bonded to additional protons that are also affected in the V13T mutant.

Nonetheless, the RMSD values are slightly higher for heme III, which is not unexpected given its proximity to the mutated residue. The three heme groups are part of the protein's hydrophobic core in which heme III is nearly perpendicular to both hemes I and IV (Figure 5.1). Therefore, the nearest heme substituents of these two hemes are affected by the heme III ring-current effects and their chemical shifts are expected to be perturbed by minor structural rearrangements on this heme. For this reason, the RMSD values for hemes I and IV are comparable (Figure 5.3A). Also, considerable chemical shift variations are observed for heme I substituents 8^2CH_3, 8^1H and 10H even though this heme is located the farthest from V13 residue.

As mentioned above, two conformations within the heme core are observable for V13I, V13S and V13T mutants (Figure 5.2). The conformations showing the lowest and the highest chemical shift differences compared to PpcA were designated by 'conformation A' and 'conformation B', respectively. The data obtained shows that the perturbation on the heme core increases in the following order: V13I (total RMSD value of 0.06 ppm), V13S (0.11 ppm) and V13T (0.13 ppm), suggesting that the volume and polarity of the side chain of residue 13 is crucial to maintain a single conformation of the protein (Figure 5.3B). In fact, in the V13A mutant, the small volume of the alanine residue (92 $Å^3$) [11, 12] compared to valine (142 $Å^3$) allows this residue to fit in the hydrophobic core and only one conformation is observed in solution. On the other hand, the larger volume of isoleucine (169 $Å^3$) perturbs the protein's hydrophobic core yielding two conformations in solution. However, the perturbation of the heme core is even higher for the polar mutants. In this case, and despite their smaller volume compared to valine (99 and 122 $Å^3$ for V13S and V13T, respectively) their polarity has a significant impact on the cytochrome's heme core conformation. In both cases, two conformations were also observed in solution, which differ the most for the larger residue (V13T). For the three mutants displaying two different conformations, the most affected heme substituents are conserved and are located in the edges pointing inside the protein hydrophobic core (*cf.* carbon atoms highlighted in red sticks in Figure 5.3C). Compared to V13I and V13S, three additional heme protons (corresponding to carbon atoms labeled with asterisks in Figure 5.3C) are also affected in V13T which is in accordance with the largest perturbations caused by this replacement. Therefore, Val[13] occupies a pivotal position in PpcA, ensuring the maintenance of a unique heme conformation that is essential for an efficient electron transfer with the redox partners.

5.1.2 Structural and Functional Role of Phe[6] and Trp[45]

5.1.2.1 Cloning, Overexpression and Purification of the Mutants

As described in section 2.1, given the high GC content of the primers harboring the mutations, two PCR mix reactions were tested in order to select the most favorable conditions: one using Phusion High Fidelity Buffer and other using Phusion GC Buffer with the addition of DMSO. The 0.8% agarose gel of the PCR products of the tested PCR mix reactions is shown in Figure

5.4. Lanes 1 and 3, which correspond to the utilization of GC Buffer and DMSO in the PCR mix reactions, present the most intense bands and the PCR products possess the correct size – 4 461 bp – therefore establishing the utilization of these components for the mutagenesis protocol. The mutations were confirmed by DNA sequencing and the plasmids containing the desired mutations were used for protein overexpression.

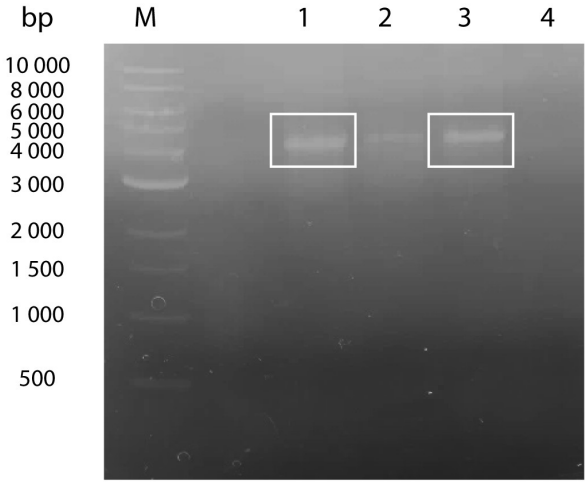


Figure 5.4 – Agarose gel electrophoresis of the PCR products for the assessment of the best PCR conditions. The gel contained 0.8% of agarose in 1x TAE (Tris-acetate EDTA) buffer. Gel lanes: M – DNA ladder; 1 – PCR product of W45M mutant in GC Buffer with the addition of DMSO; 2 – PCR product of W45M mutant in Phusion High Fidelity Buffer without the addition of DMSO; 3 – PCR product of F6L mutant in GC Buffer with the addition of DMSO; 4 – PCR product of F6L mutant in Phusion High Fidelity Buffer without the addition of DMSO. The more intense bands, in lanes 1 and 3, are highlighted with a white rectangle.

The replacement of a phenylalanine residue by a leucine (F6L) and the replacement of a tryptophan by a methionine (W45M) does not affect the protein's pI (9.32 [6]) since the side chains are uncharged. Therefore, the protein expression and purification protocol was the same used for the wild-type cytochrome [6]. For simplicity, only the chromatogram regarding the purification of ^{15}N-labeled F6L is presented in Figure 5.5, whereas the remaining chromatograms are presented in the Appendix A (Figures A1.6 to A1.8). The purification of the mutants encompassed a first purification step of cation exchange chromatography (Figure 5.5A). Given the very basic pI of these proteins, they strongly bind a cation exchange column equilibrated with 10 mM Tris-HCl pH 8 buffer and are eluted with a sodium chloride gradient at 89% of buffer (which corresponds to 267 mM NaCl). The identification of the peak corresponding to PpcA is facilitated by its intense red color. Even though this step eliminates some contaminant proteins, it must be complemented with a polishing step of size exclusion chromatography (*cf.* Figure 5.6 lanes 2 and 3). F6L and W45M mutants' theoretical molecular weight is 9 786 and 9 766 Da (determined by the Compute pI/MW Tool in ExPASy [13] to which was added the heme *c* molecular weight of 618 Da), respectively and, therefore, were

eluted at approximately the same volume – 90 mL – with 100 mM sodium phosphate buffer pH 8 (Figure 5.5B).

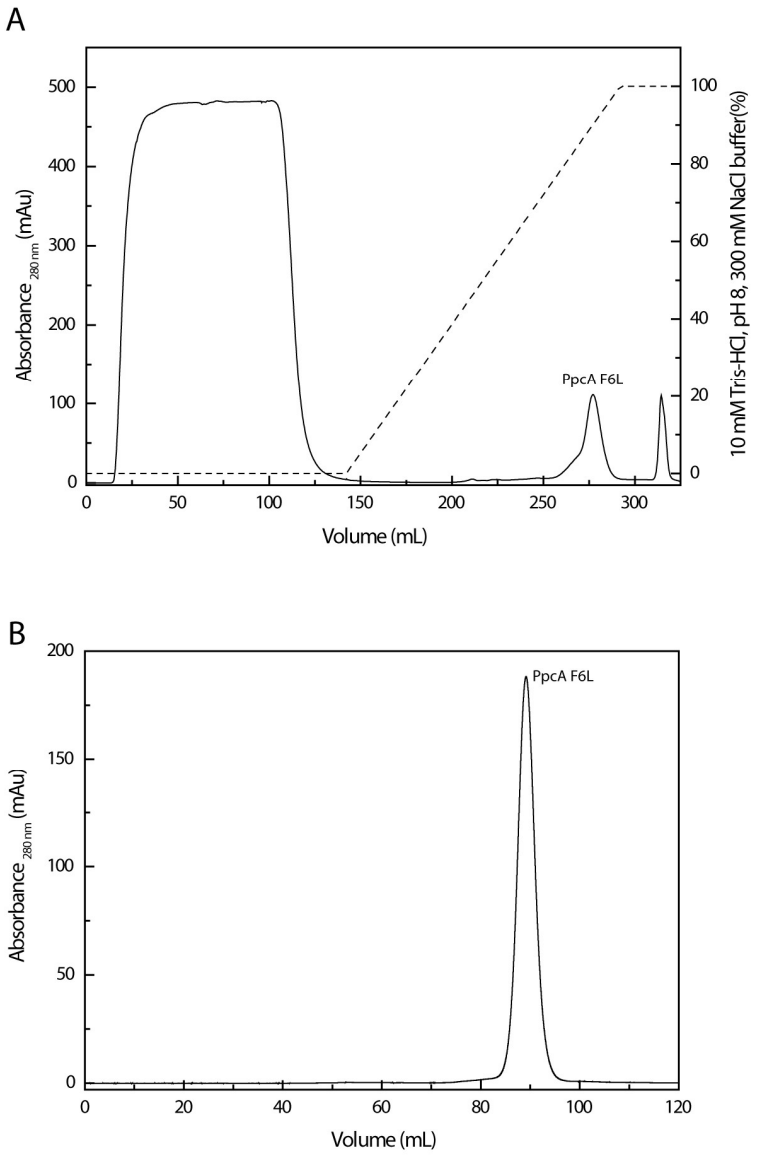


Figure 5.5 - Purification of ^{15}N labeled PpcA F6L. (A) Chromatogram correspondent to the cation exchange chromatography step. The column was equilibrated with 10 mM Tris-HCl buffer, pH 8 and the protein was eluted with a flow of 1 mL/min with a 150 mL gradient of 10 mM Tris-HCl buffer, pH 8, 300 mM NaCl. The solid line represents the elution profile whereas the dashed line represents the percentage of the NaCl gradient. (B) Chromatogram correspondent to the size exclusion chromatography step. The column was equilibrated with 100 mM sodium phosphate buffer pH 8 and the protein was eluted with a 0.5 mL/min flow.

The protein purity was evaluated throughout the purification steps by SDS-PAGE gel and, after the final purification step, a single band was observed in the expected molecular weight region (~10 kDa) for all mutants (Figure 5.6).

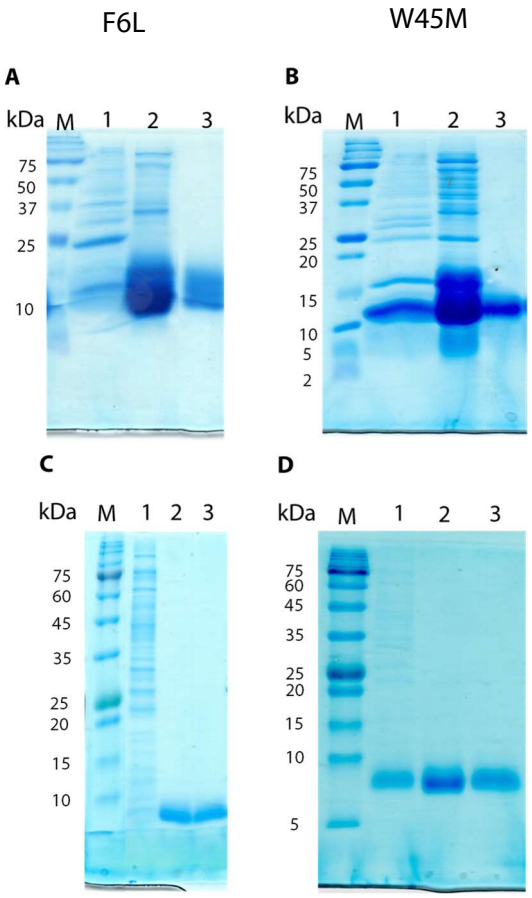


Figure 5.6 - SDS-PAGE gels of the purification process of the mutants. (A) and (C) concern the purification process of natural abundance and ^{15}N-isotopically labeled F6L, respectively, and (B) and (D) correspond to the purification process of natural abundance and ^{15}N-isotopically labeled W45M, respectively. For all gels, the M, 1, 2 and 3 lanes indicate, respectively, the protein marker, periplasmic fraction after dialysis, fraction after cation exchange chromatography and fraction after size exclusion chromatography. Protein Marker Precision Plus Protein™ Dual Xtra Standards (BioRad) was used in (A) and (B) and NZYColour Protein Marker I (NZYTech) was used in (C) and (D).

The protein yields per liter of cell culture are indicated in Table 5.1 and were determined using the PpcA molar extinction coefficient of the α band in the reduced form (ε_{552} = 118 $mM^{-1}cm^{-1}$ [6]).

Table 5.1 - Protein yields of PpcA F6L and W45M mutants expressed in natural abundance and with ^{15}N-isotopic labeling. The protein yields are referred to liter of cell culture.

Protein	Labeling	Yield (mg/L)
PpcA F6L	Natural Abundance	2.6
	^{15}N-labeling	0.8
PpcA W45M	Natural Abundance	1.4
	^{15}N-labeling	0.6

5.1.2.2 Structural Impact of the Mutations

When assessing the functional impact of mutations in a protein, it is important to ensure that the mutation does not affect the overall protein folding, otherwise the conclusions drawn for functional properties are biased. In cytochromes, both the polypeptide chain and the heme core provide a complementary analysis: while alterations of backbone amide signals in the 2D $^{1}H,^{15}N$-HSQC spectrum indicate a modified conformation of the polypeptide chain, the analysis of the heme substituents' signals provides valuable information regarding the hemes' geometry.

In the wild-type protein, Phe[6] and Trp[45] side chains are located in the hydrophobic core in two different structural regions (Figure 5.7): Phe[6] belongs to the first strand of the β-sheet whereas Trp[45] is part of the second α-helix that extends between residues Lys[43] and Glu[55].

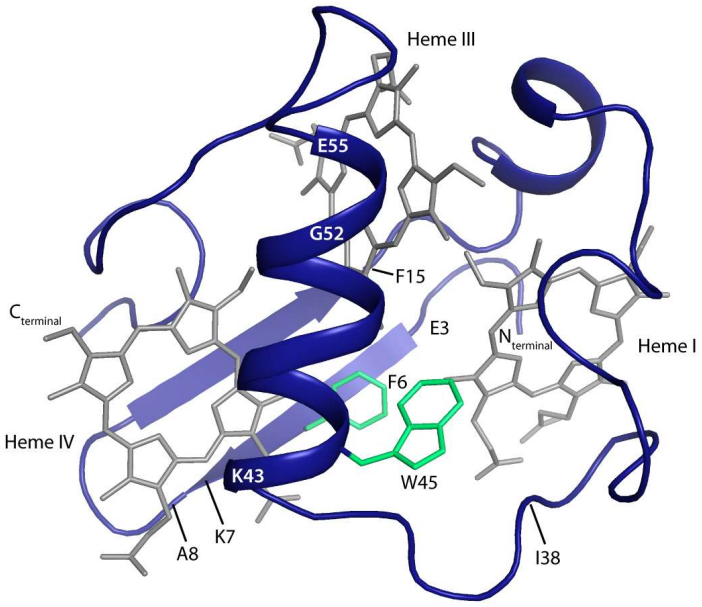


Figure 5.7 – Spatial location of the mutated residues in the solution structure of PpcA from *Gm* in the reduced state. The backbone C_α atoms are colored blue, hemes are colored gray and the Phe[6] and Trp[45] residues' side chains are represented by green sticks. Other structurally important amino acids, under the scope of this study, are labeled in the structure. Figure was generated with PyMOL [9].

The analysis of the amide signals in the 2D $^{1}H,^{15}N$-HSQC spectrum of both mutants (Figure 5.8) clearly indicates their different structural location. In F6L, the differences are essentially confined to the β-strand segment of the protein that includes residues Glu[3]-Ala[8] whereas in W45M the differences are more evident between residues Ile[38] and Gly[52] that belong to the α-helix and the random coil segment immediately before. In the wild-type protein, the Phe[6] ring establishes NOE connectivities with Leu[4], Val[13] and Phe[15] residues and the Trp[45] ring is spatially close to the residues in the referred random coil segment, hence explaining the observed differences upon introduction of the mutation in these positions.

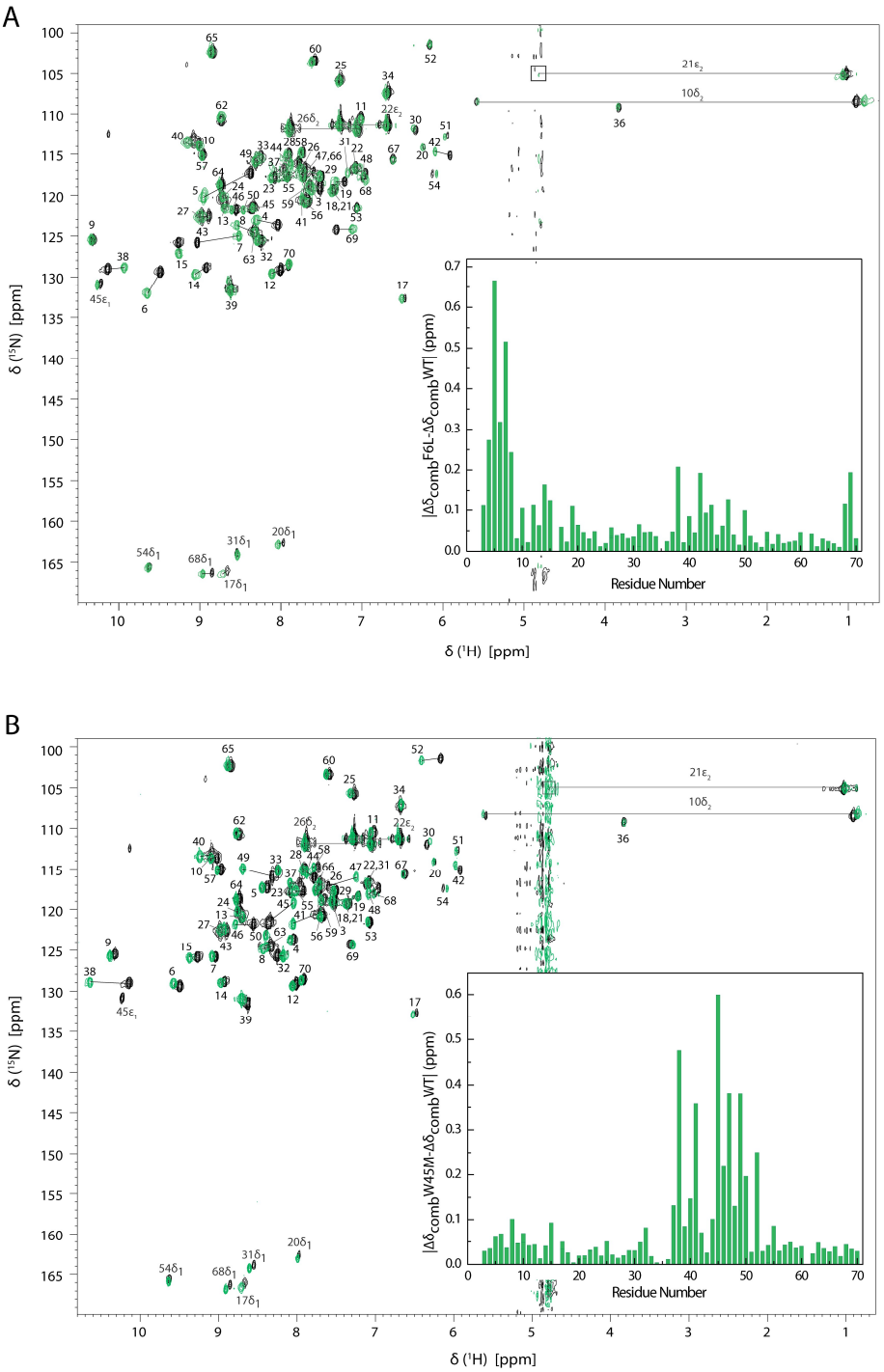


Figure 5.8 - Comparison between the $^{1}H,^{15}N$-HSQC spectrum of fully reduced F6L (A) and W45M (B) mutants (green contours) with the spectrum of wild-type PpcA (black contours). The most affected signals are connected by a straight line. Each inset represents the differences between the combined ^{1}H and ^{15}N chemical shifts observed for the mutants ($\Delta\delta_{comb}$F6L and $\Delta\delta_{comb}$W45M) and wild-type protein ($\Delta\delta_{comb}$WT). The differences were calculated using the equation $\Delta\delta_{comb}=[(\Delta\delta H)^2+ w_i (\Delta\delta N)^2]^{1/2}$, where $\Delta\delta H$ is the difference between ^{1}H shifts, $\Delta\delta N$ the difference between ^{15}N shifts, and $w_i =|\gamma\ ^{15}N\ |/|\gamma\ ^{1}H|$ is a weighting factor that accounts for the differences in nuclei sensitivity [14]. The chemical shifts are reported in Table A2.7 and Table A2.8 in the Appendix A.

The heme substituent signals were assigned using the 2D $^{1}H,^{1}H$-TOCSY and 2D $^{1}H,^{1}H$-NOESY spectra (Figure 5.9) and the presence of six pairs of signals between the thioether methines ($3^{1}H$

or 8^1H) and the thioether methyl groups (3^2CH_3 and 8^2CH_3) indicates that the heme core maintains a single conformation despite the introduction of mutated residues. The observed inter-heme NOE connectivities are almost identical to those of the wild-type protein suggesting conservation of the heme core architecture (*cf.* Figure 5.10 and Figure A1.5 in the Appendix A).

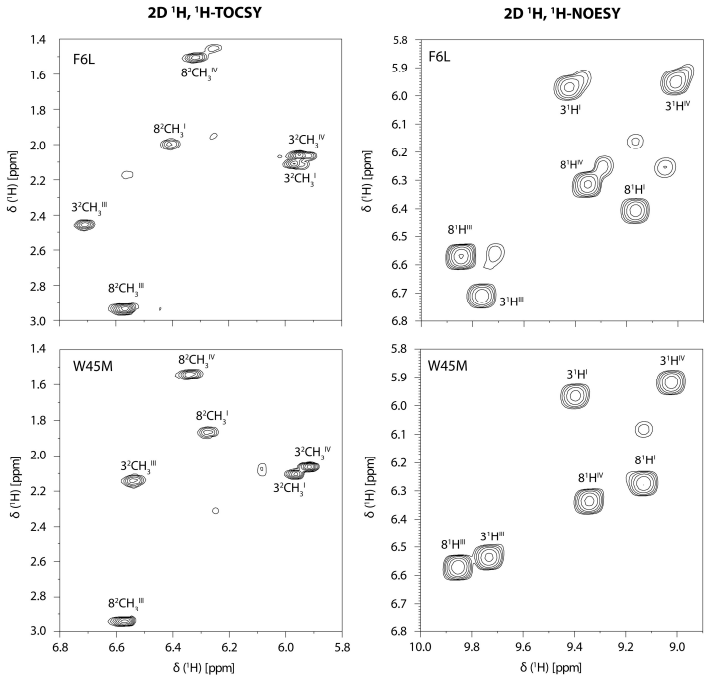


Figure 5.9 - Selected regions of 2D ^{1}H,^{1}H-TOCSY and 2D ^{1}H,^{1}H-NOESY NMR spectra to illustrate the observed differences in the hemes' substituents. The connectivities between the thioether methines (3^1H or 8^1H) and the thioether methyl groups (3^2CH_3 and 8^2CH_3) are illustrated in the 2D ^{1}H,^{1}H-TOCSY spectra (left panels). The NOE connectivities between 5H and 10H meso protons and their respective thioether methine protons are illustrated in the 2D ^{1}H,^{1}H-NOESY spectra (right panels).

Within the framework of a conserved heme core, as confirmed by the analysis of the inter-heme NOE connectivities, the chemical shift differences observed between the heme signals of the mutants and the wild-type protein (Figure 5.10) are explained by the different chemical environment provided by the substituting amino acids and, more importantly, by the absence of the aromatic ring-current contribution from the side chains of Phe[6] and Trp[45] to the observed chemical shifts. It is worth noting that, for both mutants, the differences are of the same magnitude (with exception of $3^2CH_3^{III}$ substituent in W45M) which indicates that the mutations have a comparable impact in the heme substituent chemical shifts. However, the chemical shifts of the heme substituents are differently affected as a result of the different spatial localization of the aromatic residues in the protein's hydrophobic core. In the wild-type cytochrome, the side chain of Phe[6] is almost parallel to hemes I and IV and the substitution by a leucine, which possesses a smaller volume (168 Å^3 of leucine *versus* 203 Å^3 of phenylalanine) only affected the local conformation in the β-sheet segment. The most affected substituent is $12^1CH_3^I$ followed by $12^1CH_3^{IV}$ and $2^1CH_3^{III}$, and then by $20H^{III}$ and $10H^I$ (Figure 5.10A). These are all facing the hydrophobic core of the protein and establish NOE connectivities in the wild-type protein either

directly with Phe[6] (such as $12^1CH_3^I$) or with residues of the β-sheet such as Val[13] and Phe[15]. Therefore, upon substitution of this residue, the observed alterations are expected.

In W45M, the majority of the affected heme substituents are consistent with the location of Trp[45] between hemes I and III (Figure 5.10C). The most affected substituent is $3^2CH_3^{III}$ which is nearest the ring, followed by $10H^I$, 3^1H^{III}, 8^1H^I, $12^1CH_3^I$, $12^1CH_3^{IV}$ and $8^2CH_3^I$. In a lesser extent, the heme substituents $15H^I$, $2^1CH_3^{III}$ and $7^1CH_3^{III}$ are also affected. Similarly to what occurs in the F6L mutant, all the substituents are facing the protein's hydrophobic core. The substitution of a bulky amino acid by one with a lesser volume (240 Å^3 in tryptophan *versus* 171 Å^3 in methionine) most likely impacts the chemical environment in the hydrophobic core. Heme I and III substituents are the most affected, which is expected given their closer contact with the Trp[45] ring in the wild-type protein.

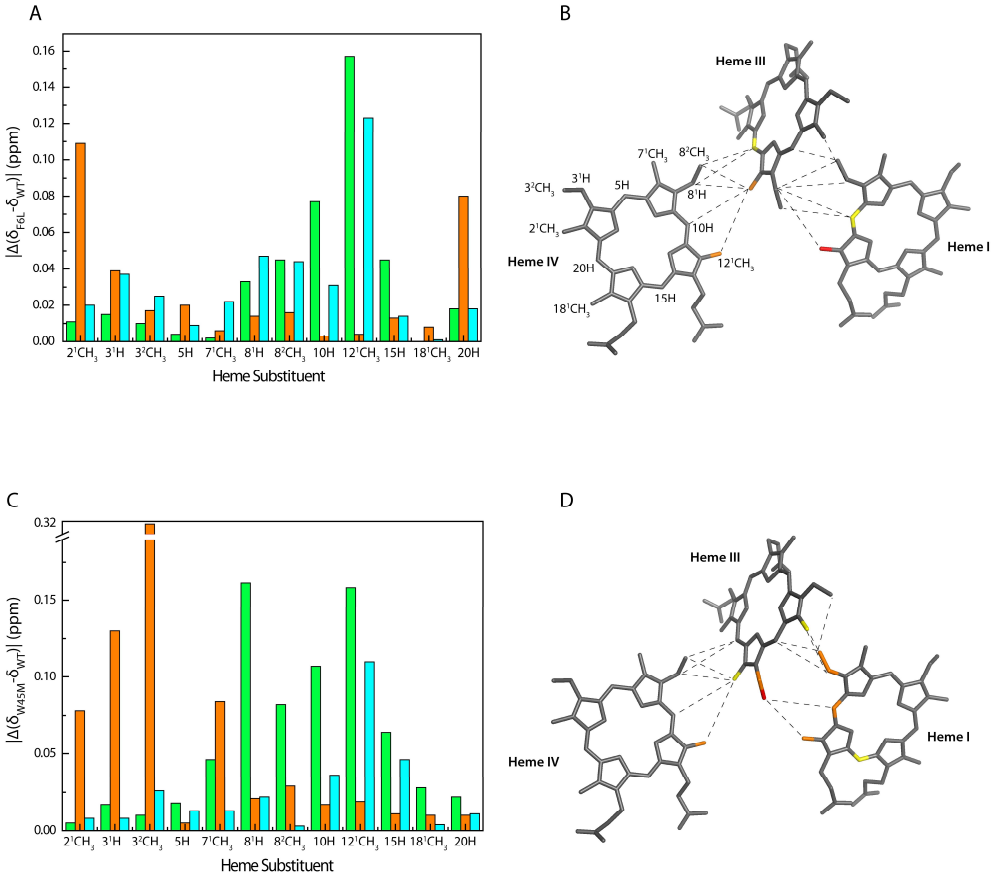


Figure 5.10 - Impact of the F6L (A and B panels) and W45M (C and D panels) mutations on the heme core architecture. Panels A and C depict the variation of the heme proton chemical shifts between the mutants' heme substituents (δ_{F6L} or δ_{W45M}) and those of the wild-type protein (δ_{WT}). The green, orange, and blue bars correspond to hemes I, III, and IV, respectively. Panels B and D indicate the structural location of the most affected substituents and the unambiguous inter-heme NOE connectivities found. For F6L (panel B), the carbon atoms bonded to the affected protons are colored accordingly to the extent to which the substituents are affected: red (Δ>0.13 ppm); orange (0.13 ppm>Δ>0.09 ppm) and yellow (0.09 ppm>Δ>0.06 ppm). For W45M (panel D), the carbon atoms bonded to the affected protons are also colored in accordance to the extent to which the substituents are affected: red (Δ=0.28 ppm); orange (0.28 ppm>Δ>0.08 ppm) and yellow (0.08 ppm>Δ>0.06 ppm). The chemical shifts of the heme substituents are reported in Table A2.9 of the Appendix A.

5.1.2.3 Functional Impact of the Mutations

Residues Phe[6] and Trp[45] are conserved within the PpcA family of *Gm* but not in the family of PpcA from *Gs* and, given their location in the protein's hydrophobic core and the observed multitude of NOE connectivities they establish with the heme groups, we have hypothesized, as previously referred, that these aromatic amino acids could be partially responsible for the distinct functional properties of the two cytochromes, contributing to a more positive functional reduction potential range observed in PpcA from *Gm*.

The spectroscopic properties of cytochromes are highly advantageous for following their global redox profile by UV-Visible spectroscopy and permit the determination of the apparent reduction potential (E_{app}). This information will be used for probing the effects of the mutations in the cytochromes' reduction potential without the need to pursue a complete thermodynamic characterization. Figure 5.11 shows the spectra of oxidized and reduced F6L mutant (respective spectra of W45M mutant is shown in Figure A1.9 on the Appendix A) where the characteristic absorption bands of *c*-type cytochromes are present: in the oxidized state there are bands at 408 (Soret) and 530 nm, and in the reduced state bands are observed at 417 (Soret), 522 (β-band) and 552 (α-band) nm. The α-band region is ideal for following the reduction and oxidation of PpcA since it absorbs distinctively in the oxidized and reduced states (*cf.* inset of Figure 5.11). The reduced fraction of the protein at each solution potential is determined by the normalized area of the 552 nm peak flanked by the isosbestic points at 545 and 560 nm and the fitting of the experimental data follows Equation 1 of section 2.3.

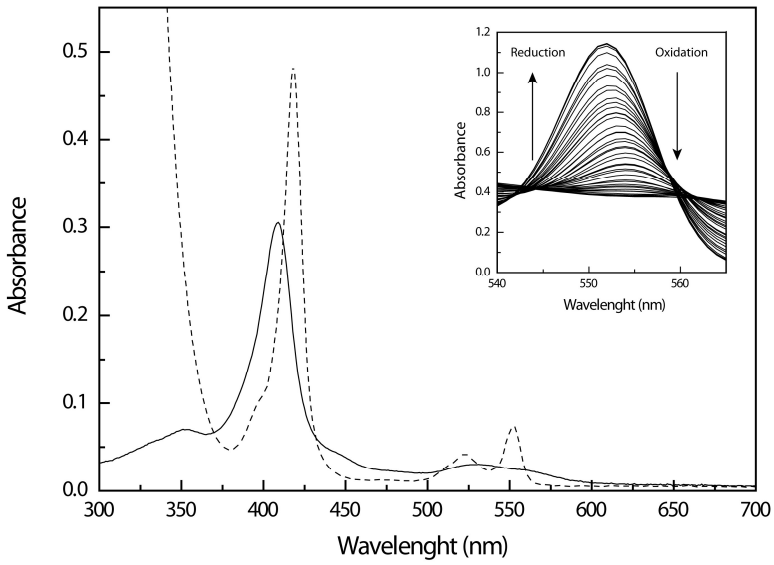


Figure 5.11 - UV-Visible spectra of PpcA F6L in the fully oxidized (solid line) and reduced (dashed line) states. The inset shows the α-band region of the visible spectra of F6L redox titration at pH 8 and 15 °C with arrows indicating the absorbance variation of reduction and oxidation.

The fitting of the UV-Visible redox titration data of F6L and W45M mutants, as well as a comparison between the fitting curves of the mutants, the wild-type protein and PpcA from *Gs* is indicated in Figure 5.12. The E_{app} value determined for F6L is of -85 mV and for W45M is of -106 mV (Table 5.2).

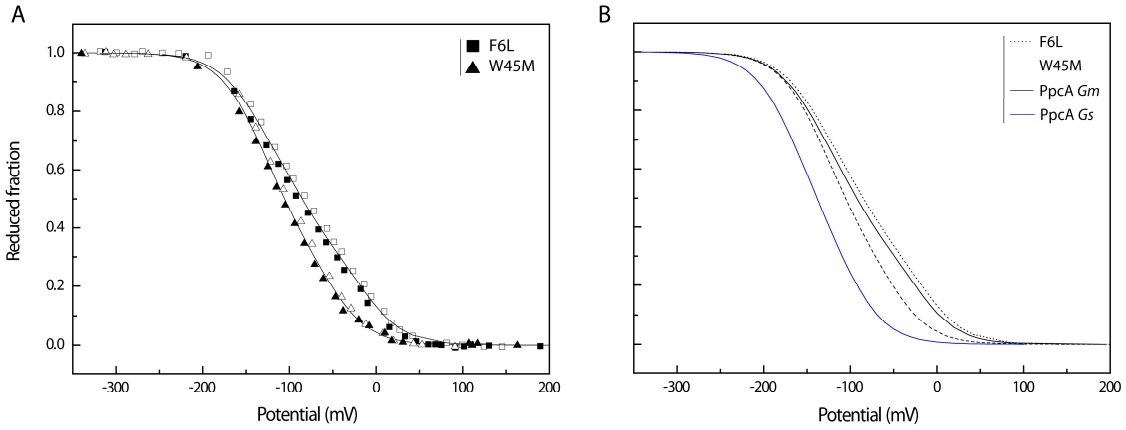


Figure 5.12 - Redox titrations followed by UV-Visible spectroscopy for PpcA F6L (squares) and W45M (triangles) mutants at pH 8 and 15 °C. (A) The solid lines indicate the results of the fits to the Equation 1 (see section 2.3). The E_1, E_2, and E_3 values obtained from the fits are indicated in Table 5.2. (B) Comparison of the fittings for the PpcA mutants, wild-type PpcA [6] and PpcA from *Gs* [10].

Table 5.2 - Apparent and macroscopic reduction potentials (*versus* NHE) for PpcA mutants F6L and W45M, wild-type PpcA (PpcA *Gm*) and PpcA from *Gs* (PpcA *Gs*) at pH 8, 15 ºC. The apparent reduction potentials (E_{app}) correspond to the point at which the oxidized and reduced fractions are equal. E_1, E_2, and E_3 are the macroscopic reduction potentials for the first, second and third oxidation steps, respectively (*cf.* Figure 1.7). Standard errors are indicated in parenthesis.

Protein	E_{app} (mV)	E_1 (mV)	E_2 (mV)	E_3 (mV)
PpcA *Gs* [10]	-138 (5)	-182 (5)	-139 (5)	-93 (5)
PpcA *Gm*[6]	-93(4)	-150 (4)	-95 (4)	-20 (4)
PpcA F6L	-85 (6)	-146 (6)	-87 (6)	-13 (6)
PpcA W45M	-106 (1)	-151 (1)	-109 (1)	-49 (1)

The heme reduction potential values are modulated by a series of factors such as the axial ligands' orientation, the heme solvent exposure, orientation and hydrogen bonding of heme propionates [15]. Based on the E_{app} values of PpcA from *Gm* and *Gs*, it was expected that the substitution of particular residues in close contact with the heme groups in PpcA from *Gm* for its counterpart in PpcA from *Gs* would lower the reduction potential. This in fact occurs for the W45M mutant in which the E_{app} value is 13 mV lower (~ 30%). The absence of a more hydrophobic residue in position 45 might increase the hemes' I and III solvent exposure and therefore explain the observed E_{app} value. On the contrary, the removal of the aromatic amino acid in position 6 led to an increase in the E_{app} value by 8 mV in comparison with the wild-type protein. The increase in the reduction potential value of F6L may be attributed to the observed

conformational changes in the β-sheet segment which may decrease hemes I and IV solvent exposure. It is worth noting that the amide backbone signals of Lys[7] and Ala[8] were considerably affected by the mutation (*cf.* Figure 5.8) and that these residues are in close proximity with the propionates of heme IV which orientation may be, therefore, altered. The observed difference in the E_{app} value is smaller compared to the verified in the W45M mutant (Figure 5.12 and Table 5.2), suggesting that Trp[45] might play a more important role in the modulation of heme reduction potential values than Phe[6]. Nevertheless, these results only assess the macroscopic impact of the mutations and therefore a detailed thermodynamic study of these mutants will shed light onto their influence on the hemes' reduction potential values, permitting the better understanding of the impact of these mutations in the heme redox properties.

5.2 Conclusions

Mutations in Phe[6], Val[13] and Trp[45] were studied in order to evaluate (i) the structural importance of Val[13] in the PpcA family from *Gs* and from *Gm* and (ii) the structural and functional impact of the F6L and W45M mutations in PpcA from *Gm*. NMR was used to probe the structural impact of the mutations and UV-Visible redox titrations were used for determining the apparent reduction potential value of F6L and W45M mutants. NMR analysis of the heme core of Val[13] mutants revealed two conformations in solution for V13I, V13S and V13T mutants, thus elucidating the crucial role of this residue in the maintenance of a single heme core conformation in PpcA. Regarding F6L and W45M mutants, the protein's overall folding was maintained and the apparent reduction potential value of W45M was more negative than the wild-type whereas the potential of F6L was slightly more positive. These results suggest that residues Phe[6] and Trp[45] play different roles in the modulation of the reduction potentials of the wild-type protein where the absence of the latter residue contributes to a more negative reduction potential value. Future studies regarding the detailed thermodynamic characterization of F6L and W45M mutants will reveal the impact of the mutations in the heme reduction potentials and studies on the impact of the other eleven non-conserved residues in PpcA from *Gm* will permit the identification of residues which are essential for the modulation of heme reduction potential values.

5.3 References

[1] J.M. Dantas, L. Morgado, Y.Y. Londer, A.P. Fernandes, R.O. Louro, P.R. Pokkuluri, M. Schiffer, C.A. Salgueiro, Pivotal role of the strictly conserved aromatic residue F15 in the cytochrome c_7 family, JBIC Journal of Biological Inorganic Chemistry, 17 (2012) 11-24.
[2] L. Morgado, J.M. Dantas, T. Simões, Y.Y. Londer, P.R. Pokkuluri, C.A. Salgueiro, Role of Met^{58} in the regulation of electron/proton transfer in trihaem cytochrome PpcA from *Geobacter sulfurreducens*, Bioscience Reports, 33 (2013) e00002.
[3] L. Morgado, S. Lourenço, Y.Y. Londer, M. Schiffer, P.R. Pokkuluri, C.A. Salgueiro, Dissecting the functional role of key residues in triheme cytochrome PpcA: a path to rational design of *G. sulfurreducens* strains with enhanced electron transfer capabilities, PLOS ONE, 9 (2014) e105566.
[4] J.M. Dantas, T. Simões, L. Morgado, C. Caciones, A.P. Fernandes, M.A. Silva, M. Bruix, P.R. Pokkuluri, C.A. Salgueiro, Unveiling the structural basis that regulates the energy transduction properties within a family of triheme cytochromes from *Geobacter sulfurreducens*, The Journal of Physical Chemistry B, 120 (2016) 10221-10233.
[5] M.R. Ferreira, J.M. Dantas, C.A. Salgueiro, The triheme cytochrome PpcF from *Geobacter metallireducens* exhibits distinct redox properties, FEBS Open Bio, 8 (2018) 1897-1910.
[6] P.C. Portela, T.M. Fernandes, J.M. Dantas, M.R. Ferreira, C.A. Salgueiro, Biochemical and functional insights on the triheme cytochrome PpcA from *Geobacter metallireducens*, Archives of Biochemistry and Biophysics, 644 (2018) 8-16.
[7] J.M. Dantas, L. Morgado, M. Aklujkar, M. Bruix, Y. Londer, M. Schiffer, P.R. Pokkuluri, C.A. Salgueiro, Rational engineering of *Geobacter sulfurreducens* electron transfer components: a foundation for building improved *Geobacter*-based bioelectrochemical technologies, Frontiers in Microbiology, 6 (2015).
[8] L. Morgado, V.B. Paixão, M. Schiffer, P.R. Pokkuluri, M. Bruix, C.A. Salgueiro, Revealing the structural origin of the redox-Bohr effect: the first solution structure of a cytochrome from *Geobacter sulfurreducens*, Biochemical Journal, 441 (2012) 179-187.
[9] Schrodinger, LLC, The PyMOL Molecular Graphics System, Version 1.8, in, 2015.
[10] L. Morgado, M. Bruix, V. Orshonsky, Y.Y. Londer, N.E.C. Duke, X. Yang, P.R. Pokkuluri, M. Schiffer, C.A. Salgueiro, Structural insights into the modulation of the redox properties of two *Geobacter sulfurreducens* homologous triheme cytochromes, Biochimica et Biophysica Acta - Bioenergetics, 1777 (2008) 1157-1165.
[11] F.M. Richards, Areas, volumes, packing, and protein structure, Annual Review of Biophysics and Bioengineering, 6 (1977) 151-176.
[12] G. Baumann, C. Froömmel, C. Sander, Polarity as a criterion in protein design, Protein Engineering, Design and Selection, 2 (1989) 329-334.
[13] E. Gasteiger, A. Gattiker, C. Hoogland, I. Ivanyi, R.D. Appel, A. Bairoch, ExPASy: the proteomics server for in-depth protein knowledge and analysis, Nucleic Acids Research, 31 (2003) 3784-3788.
[14] F.H. Schumann, H. Riepl, T. Maurer, W. Gronwald, K.-P. Neidig, H.R. Kalbitzer, Combined chemical shift changes and amino acid specific chemical shift mapping of protein–protein interactions, Journal of Biomolecular NMR, 39 (2007) 275-289.
[15] A. Dolla, L. Blanchard, F. Guerlesquin, M. Bruschi, The protein moiety modulates the redox potential in cytochromes *c*, Biochimie, 76 (1994) 471-479.

Chapter 6
Final Conclusions and Future Work

6 Final Conclusions and Future Work

Multiheme cytochromes are essential players in the extracellular electron transfer mechanisms in *Geobacter* bacteria. PpcA is a highly abundant periplasmic triheme cytochrome both in *G. sulfurreducens* (*Gs*) and *G. metallireducens* (*Gm*) and has been proven to be an important electron transfer component in these mechanisms. The high identity of PpcA from *Gm* and *Gs* (80%) but their markedly different functional properties set the stage to study PpcA from *Gm* with the ultimate goal of identifying key residues responsible for the modulation of the redox properties in these cytochromes.

The work developed in this thesis had as a first goal the structural and functional characterization of PpcA from *Gm* in the reduced and oxidized states. The determination of the preliminary NMR solution structure in the reduced state has revealed important localized structural differences compared to its homologue in *Gs*, especially in the polypeptide segments near hemes I and III. In the oxidized state, the backbone, side chain and heme signal's assignment also corroborated these observations. The pH titrations followed by NMR in the reduced and oxidized states have attributed the redox-Bohr center to the heme IV propionate 13, which is in accordance with the previous hypothesis proposed from the thermodynamic studies of PpcA from *Gm*. Moreover, the residues which chemical shifts were most affected by the pH in PpcA from *Gm* indicate that the structural origin for pH-linked conformational changes in both redox states presents subtle differences when compared to PpcA from *Gs*.

The second goal of this thesis was the assessment of the structural and functional importance of one conserved and two non-conserved residues in the PpcA families of *Gm* and *Gs*. The conserved residue Val^{13} was found to be essential in the maintenance of a single heme core conformation, thus ensuring optimal electron transfer to physiological redox partners. The non-conserved residues Phe^{6} and Trp^{45} were mutated to the correspondent amino acids in PpcA from *Gs* (leucine and methionine, respectively) and were shown to exert different modulation of the reduction potentials: while mutation in the tryptophan brought closer the reduction potential of PpcA from *Gm* to that of PpcA from *Gs*, the mutation in phenylalanine had the opposite effect, although at a smaller extent.

The results presented in this thesis open many possible future studies. In the reduced state, the conclusion of the solution structure will allow a thorough structural comparison with PpcA from *Gs*, thus elucidating the observed differences in this work. In the oxidized state, the NMR assignments establish foundations not only for the future determination of the solution structure in this state but also for the mapping of interactions with physiological redox partners which will permit the better understanding of extracellular electron transfer pathways in *Gm*. Finally, site-directed mutagenesis on the remaining eleven non-conserved amino acids in PpcA from *Gm* (compared to PpcA from *Gs*) and the assessment of the structural and functional impact of these mutations will permit the identification of key residues responsible for the modulation of redox potentials and contribute to the rational design of proteins with specific electron transfer properties.

A
Appendix

A.1 Supplementary Figures

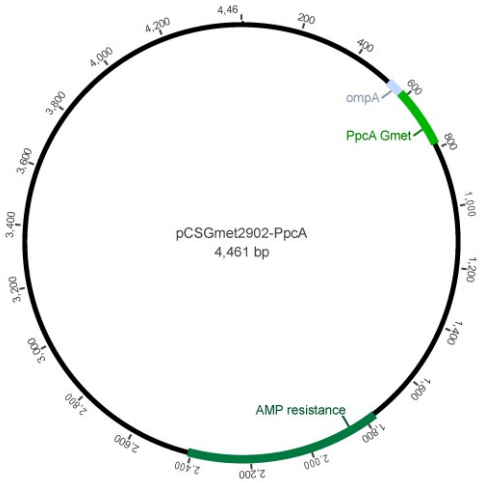


Figure A1.1 - pCSGmet2902 plasmid. The plasmid possesses an ampicillin resistance marker and contains the ompA signal sequence and the gene of PpcA from *Gm* (Gmet2902) both under the control of the *lac* promoter.

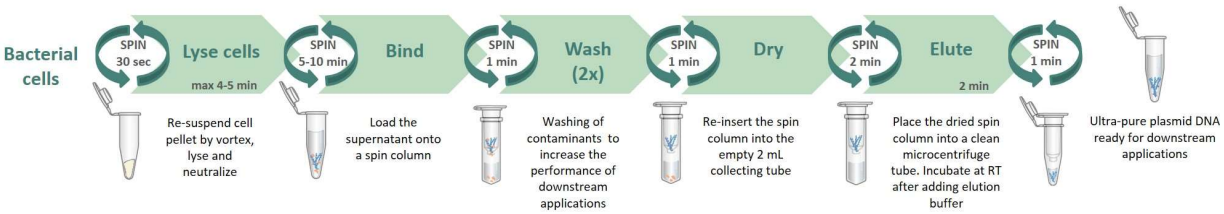


Figure A1.2 - Overview of the NZYMiniprep protocol used for the isolation of plasmid DNA. Figure retrieved from the kit's instructions.

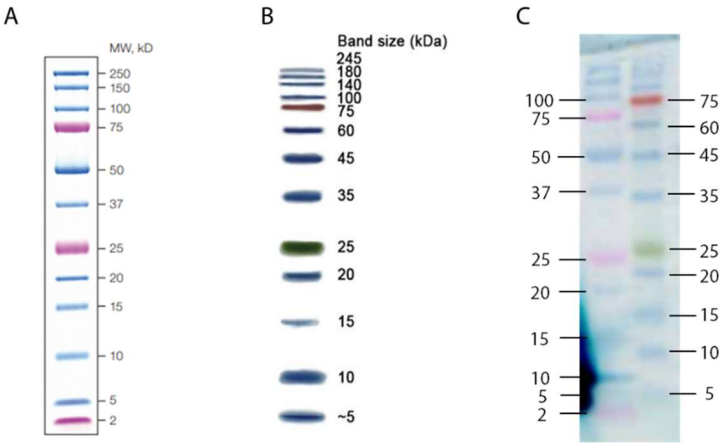


Figure A1.3 - Protein molecular weight markers used in this work for the SDS-PAGE gels. (A) Protein Marker Precision Plus Protein™ Dual Xtra Standards (BioRad). (B) NZYColour Protein Marker I (NZYTech). (C) Comparison of the migration of the two protein markers in a 15% SDS-PAGE gel. The numbers indicate the molecular weight (in kDa) of each band.

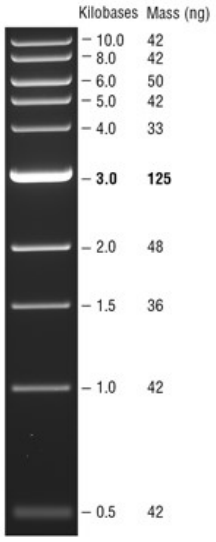


Figure A1. 4 - DNA ladder used in this work for the 0.8% agarose gels. 1 kb DNA ladder (New England Biolabs). The numbers indicate the number of kilobases (kb) and the mass (in ng) of each band.

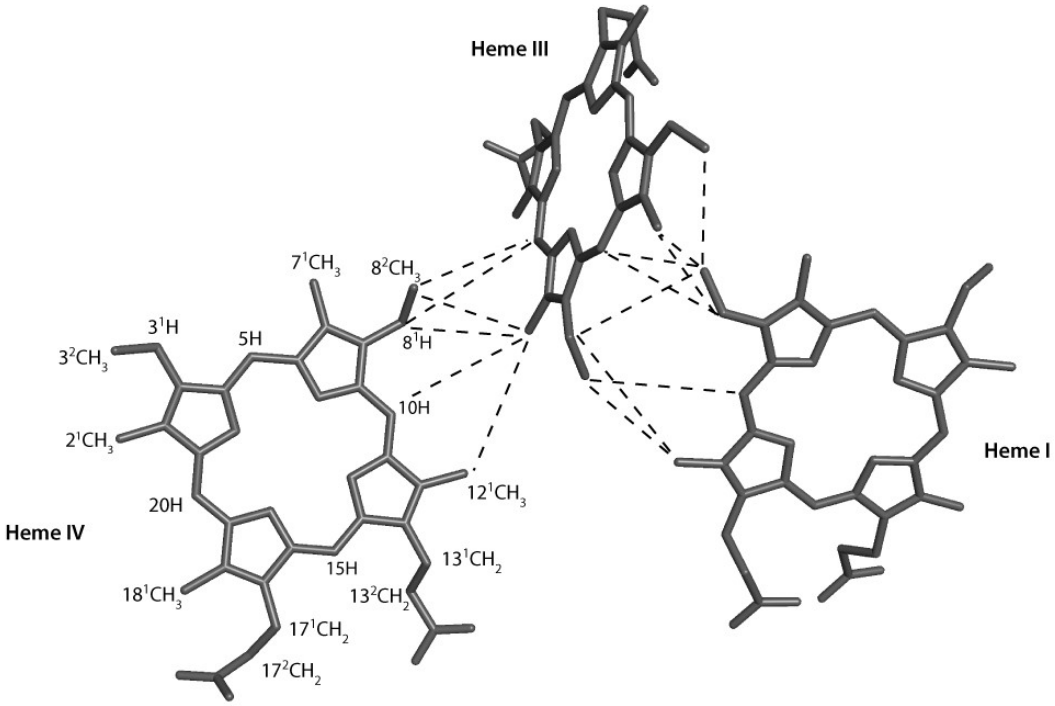


Figure A1. 5 - Inter-heme NOE connectivities for PpcA from *Gm* in the reduced state. The represented heme core is the one of the lowest-energy solution structure of PpcA in the reduced state. Black dashed lines indicate the inter-heme NOE connectivities observed in the 2D ^{1}H,^{1}H-NOESY spectrum of PpcA. The IUPAC-IUB nomenclature for tetrapyrroles is represented on heme IV.

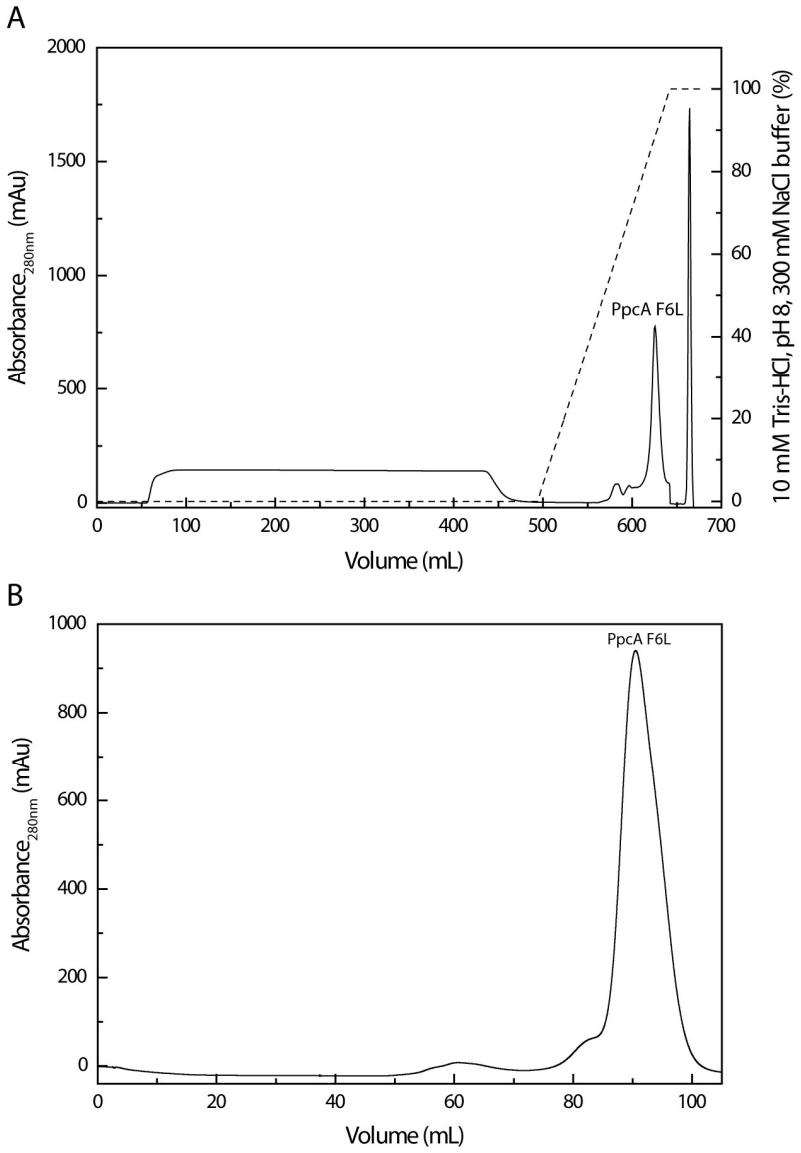


Figure A1. 6 - Purification of natural abundance PpcA F6L. (A) Chromatogram of the first purification step by cation exchange chromatography. The column was equilibrated with 10 mM Tris-HCl buffer, pH 8 and the protein was eluted with a flow of 1 mL/min with a 150 mL gradient of 10 mM Tris-HCl buffer, pH 8, 300 mM NaCl. The solid line represents the elution profile whereas the dashed line represents the NaCl gradient. (B) Chromatogram of the second purification step by size exclusion chromatography. The column was equilibrated with 100 mM sodium phosphate buffer pH 8 and the protein was eluted with a 0.5 mL/min flow.

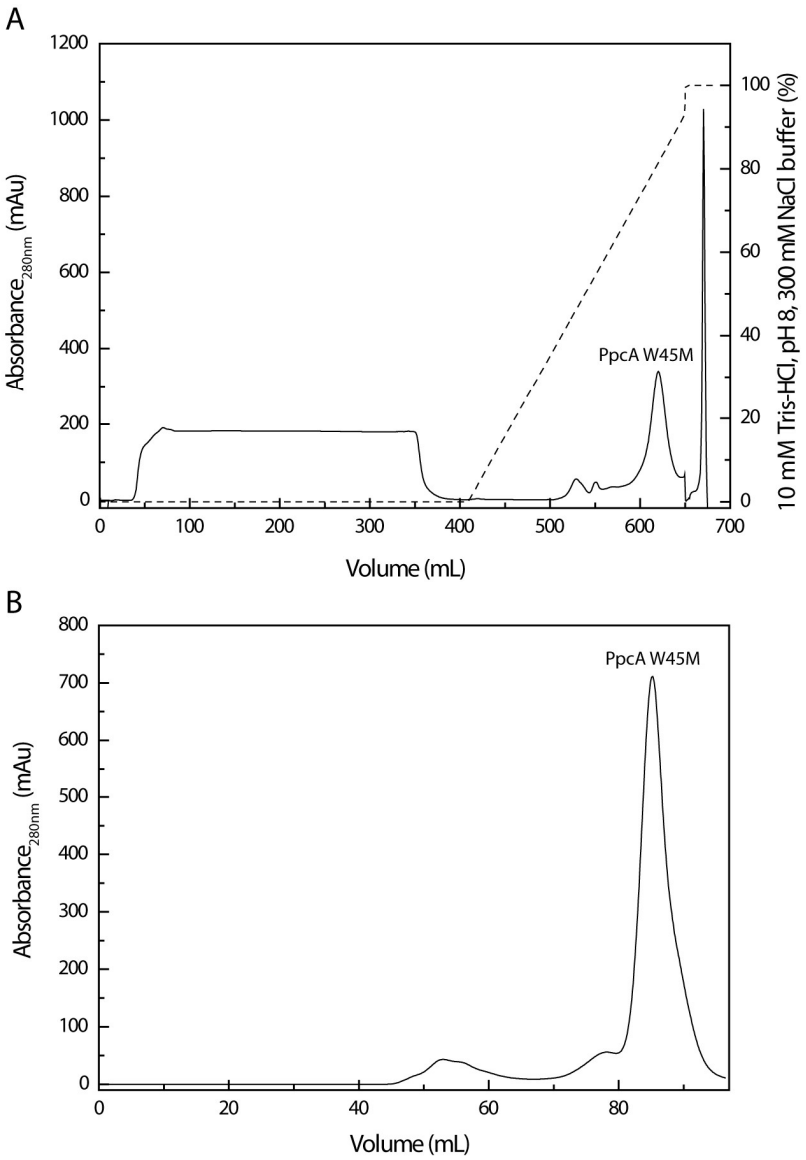


Figure A1. 7 - Purification of natural abundance PpcA W45M. (A) Chromatogram of the first purification step by cation exchange chromatography. The column was equilibrated with 10 mM Tris-HCl buffer, pH 8 and the protein was eluted with a flow of 1 mL/min with a 150 mL gradient of 10 mM Tris-HCl buffer, pH 8, 300 mM NaCl. The solid line represents the elution profile whereas the dashed line represents the NaCl gradient. (B) Chromatogram of the second purification step by size exclusion chromatography. The column was equilibrated with 100 mM sodium phosphate buffer pH 8 and the protein was eluted with a 0.5 mL/min flow.

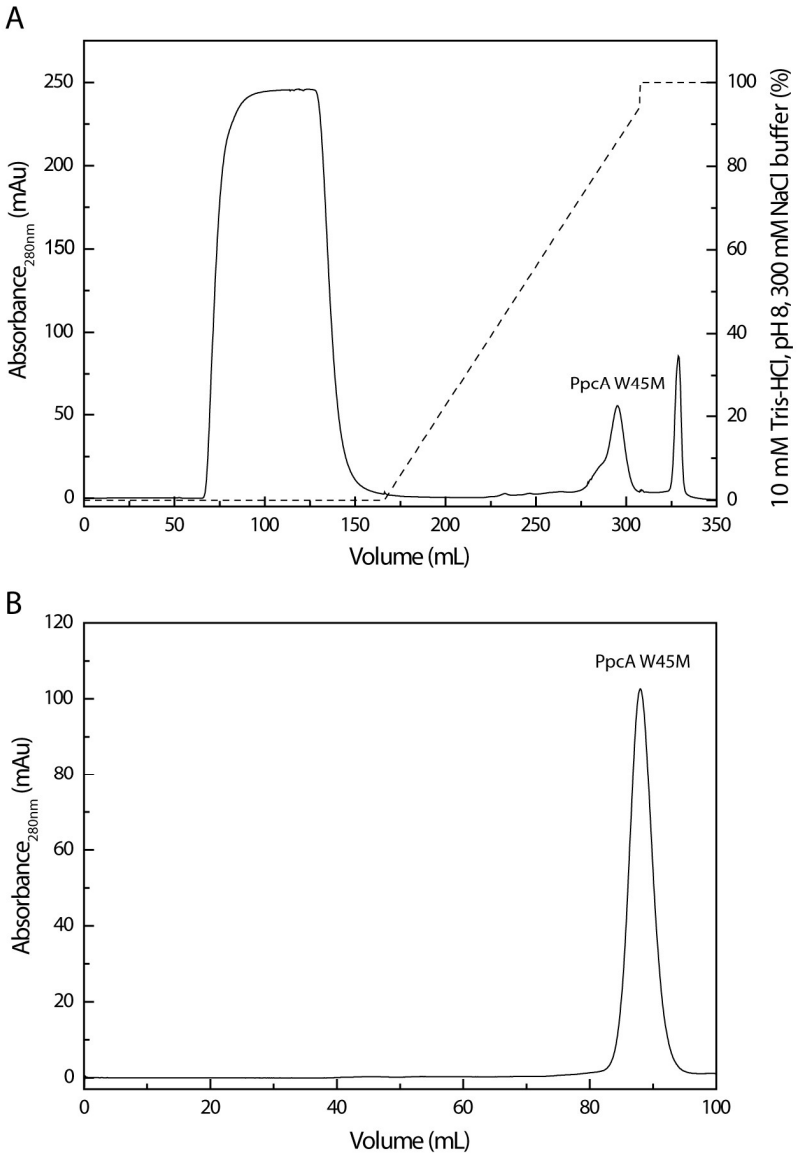


Figure A1. 8 - Purification of ^{15}N labeled PpcA W45M. (A) Chromatogram of the first purification step by cation exchange chromatography. The column was equilibrated with 10 mM Tris-HCl buffer, pH 8 and the protein was eluted with a flow of 1 mL/min with a 150 mL gradient of 10 mM Tris-HCl buffer, pH 8, 300 mM NaCl. The solid line represents the elution profile whereas the dashed line represents the NaCl gradient. (B) Chromatogram of the second purification step by size exclusion chromatography. The column was equilibrated with 100 mM sodium phosphate buffer pH 8 and the protein was eluted with a 0.5 mL/min flow.

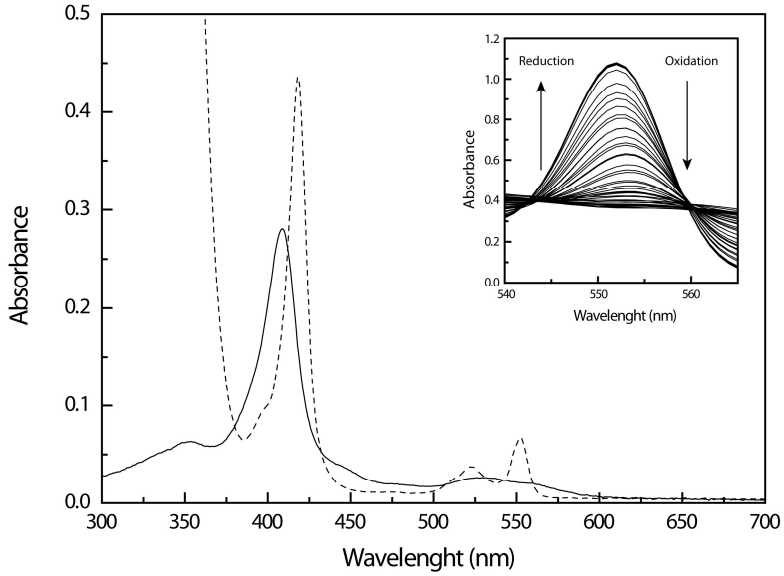


Figure A1. 9 - UV-Visible spectra of PpcA W45M in the fully oxidized (solid line) and reduced (dashed line) states. The inset shows the α-band region of the visible spectra of W45M redox titration at pH 8 and 15 °C with arrows indicating the absorbance variation of the reduction and oxidation directions.

A.2 Supplementary Tables

Table A2. 1 - Genotype of *E. coli* DH5α cells.

F⁻ φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F)U169 *rec*A1 *end*A1 *hsd*R17(r_K^-, m_K^+) *pho*A *sup*E44 λ⁻ *thi*-1 *gyr*A96 *rel*A1	
Allele	**Description**
F⁻	Does not carry the F plasmid responsible for conjugation
φ80*lac*ZΔM15	Cell carries the lambdoid prophage φ80 with the partial deletion of the lacZ gene
Δ(*lac*ZYA-*arg*F)U169	High level of resistance to hydrogen peroxide
*rec*A1	Mutation in the DNA repair system for reduced occurrence of unwanted recombination in cloned DNA
*end*A1	Lack of endonuclease I
*hsd*R17(r_K^-, m_K^+)	Mutation in the methylation system, preventing cleavage of heterologous DNA
*pho*A	Activity of alkaline phosphatase is abolished
*sup*E44	Suppression of UAG codons by insertion of glutamine, required for phage growth
λ⁻	Deletion of lambda lysogen
thi-1	Incapability of synthesizing thiamine
*gyr*A96	Mutation in DNA gyrase which conveys nalidixic acid resistance
*rel*A1	Accumulation of neutral phospholipids and saturated fatty acids that increase the fragility of the membrane upon sonication or osmotic shock

Table A2. 2 - Genotype of *E. coli* BL21(DE3) cells.

F⁻*ompT hsd* S_B(r_B^- m_B^-)*gal dcm* (DE3)	
Allele	**Description**
F⁻	Does not carry the F plasmid responsible for conjugation
ompT	Mutation in outer membrane protein protease VII, reducing proteolysis of expressed proteins
hsd S_B(r_B^- m_B^-)	Mutation in the methylation system which enables the recognition of exogenous DNA
gal	Galactose non-utilizing
dcm	Mutation in the adenine and cytosine methylation system enhancing the DNA susceptibility to cleavage in the CC(A/T)GG DNA sequence
DE3	λ prophage carrying the T7 RNA polymerase gene under the control of the lac UV5 promotor, requiring IPTG for protein expression.

Table A2. 3 - Composition of the 15% SDS-PAGE gel.

Stock Solutions	Volume (µL)	
	Stacking Gel (5 %)	Running Gel (15 %)
1.5 M Tris-HCl pH 8.8 Buffer	-	750
0.5 M Tris-HCl pH 6.8 Buffer	450	-
40 % acrylamide/bis-acrylamide (37.5:1)	225	1 880
SDS 10%	18	50
H_2O	1 107	2 280
10% Ammonium Persulfate	13.5	38
TEMED	2	2.5

Table A2. 4 - ^{1}H and ^{13}C chemical shifts (ppm) of the heme substituents of PpcA in the oxidized state, pH 5.5, 25 °C. Unassigned resonances are labeled N.A. Heme substituents are indicated in the IUPAC-IUB nomenclature for tetrapyrroles [1].

Heme substituent	Heme I		Heme III		Heme IV	
	^{1}H	^{13}C	^{1}H	^{13}C	^{1}H	^{13}C
5H	N.A	N.A	-4.84	45.4	N.A	N.A
10H	-1.99	28.38	N.A	N.A	0.24	21.09
15H	5.94	-9.85	-1.53	40.13	N.A	N.A
20H	0.03	20.94	7.48	45.86	-1.04	27.83
$2^{1}CH_3$	19.95	-39.05	10.21	-21.07	14.40	-32.84
$7^{1}CH_3$	6.26	-14.82	16.26	-39.30	10.99	-23.44
$12^{1}CH_3$	25.01	-55.99	14.97	-26.32	15.56	-36.83
$18^{1}CH_3$	16.26	-39.30	0.95	-2.77	15.40	-37.65
$3^{1}H$	N.A	N.A	-1.37	-20.94	0.22	-8.04
$8^{1}H$	-4.16	-3.03	N.A	N.A	0.03	-8.67
$3^{2}CH_3$	1.69	70.04	-1.92	62.03	1.72	62.66
$8^{2}CH_3$	-3.98	45.14	-1.15	14.56	1.22	63.51
$13^{1}CH_2$	0.99 9.03	-12.89	14.68 22.34	-60.22	3.11 4.30	-17.27
$13^{2}CH_2$	-1.20 -0.52	87.33	-2.75 -1.09	169.54	-0.68 -0.19	93.15
$17^{1}CH_2$	-1.35 3.35	-3.08	3.07 5.23	-15.89	4.01 4.51	-16.77
$17^{2}CH_2$	-0.22 0.22	68.55	-2.31 -1.98	86.27	-0.73 -0.11	94.24

Table A2. 5 - Backbone and side chain chemical shifts (ppm) of oxidized PpcA, pH 5.5, 25 °C. Atoms are indicated by the DYANA nomenclature [2].

Residue number	Residue type	Atom	Chemical shift (ppm)
1	ALA	HA	5.34
		HB1	2.78
		HB2	2.78
		HB3	2.78
		CA	53.35
		CB	20.63
2	ASP	HA	7.53
		HB2	4.18
		HB3	4.52
		CA	59.0
		CB	42.46
3	GLU	H	9.91
		HA	6.91
		HB2	3.05
		HG2	3.10
		HG3	3.15
		CA	57.18
		CB	33.63
		CG	37.60
		N	122.52
4	LEU	H	10.13
		HA	5.61
		HB2	1.63
		HB3	2.46
		HG	4.27
		QD1	1.94
		QD2	2.57
		CA	54.35
		CB	46.67
		CG	29.48
		CD1	24.85
		CD2	26.68
		N	125.50
5	THR	H	8.91
		HA	5.23
		HB	4.25
		QG	1.56
		CA	62.34
		CG2	22.06
		N	118.02
6	PHE	H	9.58
		HA	5.13
		HB2	2.74
		HB3	3.10
		HD2	6.49
		HE2	5.82
		HZ	5.19
		CA	54.98
		CB	39.46
		CD2	130.68
		CE2	130.58
		CZ	128.21
		N	128.31
7	LYS	H	8.95
		HA	5.15
		HB2	1.85
		HB3	1.94
		HG2	1.73
		HG3	1.74
		HD2	1.89
		HE2	3.16
		HE3	3.26
		CA	56.14
		CB	32.13
		CG	24.63
		CD	28.44
		CE	41.92
		N	125.83
8	ALA	H	9.85
		HA	4.87
		QB	4.40
		CA	52.60
		CB	24.33
		N	124.71
9	LYS	H	10.09
		HA	4.45
		HB2	1.63
		HB3	1.65
		HG3	1.21
		HD2	1.32

Table A2. 5 *(cont.)*

Residue number	Residue type	Atom	Chemical shift (ppm)
9	LYS	HE2	2.48
		HE3	2.53
		CA	59.30
		CB	31.94
		CG	24.23
		CD	28.83
		CE	41.49
		N	125.28
10	ASN	H	10.08
		HA	5.35
		HB2	2.04
		HB3	3.25
		CA	54.07
		CB	36.21
		N	114.83
11	GLY	H	9.73
		HA2	5.47
		HA3	7.13
		CA	47.81
		N	113.41
12	ASP	H	10.33
		HA	7.12
		HB2	3.55
		HB3	3.89
		CA	57.98
		CB	42.48
		N	133.36
13	VAL	H	10.71
		HA	6.65
		HB	3.15
		QG1	0.80
		QG2	5.98
		CA	62.19
		CB	36.51
		CG1	21.48
		CG2	22.35
		N	121.81
14	LYS	H	9.72
		HA	5.16
		HB2	2.16
		HB3	2.19
14	LYS	HG2	1.91
		HG3	2.03
		HD2	2.06
		HD3	2.12
		HE2	3.39
		CA	56.89
		CB	33.05
		CG	25.36
		CD	29.27
		CE	42.35
		N	129.58
15	PHE	H	9.77
		HA	5.82
		HB2	2.96
		HB3	3.47
		CA	52.70
		CB	43.16
		N	122.41
16	PRO	HA	6.90
		HB2	3.32
		HB3	4.19
		HG2	1.97
		HG3	3.57
		HD2	4.15
		HD3	4.29
		CA	64.07
		CB	31.12
		CG	28.26
		CD	52.24
17	HIS	H	11.60
		HA	8.40
		HB2	12.29
		HB3	17.70
		CA	73.59
		CB	27.58
		N	138.81
18	LYS	H	12.84
		HA	7.33
		HB2	4.11
		HG2	2.79
		HG3	2.97

Table A2. 5 *(cont.)*

Residue number	Residue type	Atom	Chemical shift (ppm)
18	LYS	HD2	2.94
		HD3	3.03
		HE2	3.90
		CA	62.94
		CB	33.51
		CG	26.14
		CD	30.01
		CE	43.09
		N	127.20
19	LYS	H	10.95
		HA	5.54
		HB2	3.33
		HB3	3.79
		HG2	2.12
		HG3	2.28
		HD2	1.69
		HD3	1.90
		HE2	2.61
		HE3	2.76
		CA	60.91
		CB	33.57
		CG	25.35
		CD	29.03
		CE	41.73
		N	121.71
20	HIS	H	11.84
		HA	7.68
		HB2	5.88
		HB3	8.16
		CA	81.38
		CB	20.36
		N	116.19
21	GLN	H	10.29
		HA	3.08
		HB2	2.02
		HB3	3.64
		HG2	0.21
		CA	59.26
		CB	27.03
		CG	33.53
		N	122.15
22	GLN	H	8.53
		HA	4.35
		HB2	2.60
		HB3	2.94
		HG2	2.78
		HG3	3.09
		HE21	7.21
		HE22	7.87
		CA	58.69
		CB	29.26
		CG	34.25
		N	117.55
		NE2	111.87
23	VAL	H	8.78
		HA	4.09
		HB	2.58
		QG1	1.01
		QG2	1.47
		CA	64.74
		CB	33.13
		CG1	21.36
		CG2	22.06
		N	115.20
24	VAL	H	8.46
		HA	3.34
		HB	1.17
		QG1	-0.31
		QG2	-0.43
		CA	63.72
		CB	30.53
		CG1	21.03
		CG2	21.95
		N	119.59
25	GLY	H	7.01
		HA2	3.57
		HA3	3.74
		CA	46.58
		N	105.82
26	ASN	H	6.68
		HA	4.55
		HB2	2.45

Table A2. 5 *(cont.)*

Residue number	Residue type	Atom	Chemical shift (ppm)
26	ASN	HB3	2.74
		HD21	6.78
		HD22	7.50
		CA	52.30
		CB	39.29
		N	115.54
		ND2	111.98
27	CYS	H	8.30
		HA	2.97
		HB2	4.12
		HB3	5.02
		CA	56.39
		CB	32.13
		N	122.19
28	LYS	H	8.44
		HA	5.69
		HB2	2.32
		HB3	2.34
		HG2	2.17
		HD2	2.18
		HE2	3.44
		CA	58.46
		CB	32.81
		CG	25.87
		CD	29.68
		CE	42.58
		N	117.71
29	LYS	H	7.26
		HA	3.78
		HB2	0.47
		HB3	0.93
		HG2	0.30
		HG3	0.47
		HD2	1.02
		HE2	2.12
		HE3	2.24
		CA	57.33
		CB	31.13
		CG	24.01
		CD	27.45
		CE	41.48
29	LYS	N	117.31
30	CYS	H	6.86
		HA	4.54
		HB2	0.49
		HB3	1.23
		CA	54.81
		CB	38.63
		N	109.74
31	HIS	H	12.01
		HA	10.23
		HB2	10.77
		HB3	13.00
		CA	83.16
		CB	23.34
		N	121.41
32	GLU	H	10.78
		HA	5.24
		HB2	2.91
		HB3	3.01
		HG2	2.97
		HG3	3.01
		CA	61.53
		CB	31.07
		CG	36.88
		N	129.80
33	LYS	H	10.33
		HA	5.36
		HB2	2.72
		HB3	2.80
		HG2	2.04
		HD2	2.19
		HE2	3.40
		CA	56.30
		CB	32.72
		CG	25.94
		CD	29.46
		CE	42.65
		N	117.36
34	GLY	H	8.99
		HA2	5.11
		HA3	5.20

Table A2. 5 *(cont.)*

Residue number	Residue type	Atom	Chemical shift (ppm)
34	GLY	CA	45.18
		N	108.82
35	PRO	HA	5.98
		HB2	-0.91
		HB3	-0.54
		HG2	1.40
		HG3	1.51
		HD2	3.71
		HD3	4.28
		CA	64.59
		CB	30.93
		CG	27.06
		CD	50.19
36	GLY	HA2	0.43
		HA3	3.62
		CA	44.03
37	LYS	H	7.37
		HA	2.79
		HB2	1.22
		HB3	1.25
		HG2	0.58
		HG3	0.80
		HD2	0.98
		HE2	2.47
		HE3	2.54
		CA	56.20
		CB	31.82
		CG	24.13
		CD	27.71
		CE	41.47
		N	118.64
38	ILE	H	9.79
		HA	5.19
		HB	1.16
		QG2	0.64
		QD1	1.28
		CA	61.78
		CB	40.80
		CG1	30.74
		CG2	15.61
		CD1	14.20
38	ILE	N	128.92
39	GLU	H	9.42
		HA	4.08
		HB2	2.19
		HB3	2.22
		HG2	2.38
		HG3	2.43
		CA	58.24
		CB	29.21
		CG	36.08
		N	132.07
40	GLY	H	9.00
		HA2	3.67
		HA3	4.07
		CA	45.06
		N	113.05
41	PHE	H	7.42
		HA	2.73
		HB2	2.03
		HB3	2.17
		HD2	5.35
		CA	60.58
		CB	38.43
		CD2	129.56
		N	120.09
42	GLY	HA2	2.08
		HA3	2.11
		CA	43.11
43	LYS	H	7.25
		HA	-0.09
		HB2	-0.75
		HB3	0.93
		HG2	4.10
		HD2	1.40
		HE2	2.95
		HE3	3.90
		CA	58.71
		CB	31.73
		CG	25.47
		CD	30.76
		CE	43.53

Table A2. 5 *(cont.)*

Residue number	Residue type	Atom	Chemical shift (ppm)
43	LYS	N	120.77
44	ASP	H	8.65
		HA	6.88
		HB2	3.22
		HB3	3.58
		CA	59.47
		CB	40.79
		N	118.75
45	TRP	H	8.89
		HA	4.10
		HB2	2.25
		HB3	3.17
		HD1	7.22
		HE1	10.44
		HE3	5.53
		HZ2	7.43
		HZ3	3.93
		HH2	4.48
		CA	62.01
		CB	31.46
		CD1	128.21
		CE3	119.82
		CZ2	115.24
		CZ3	117.60
		CH2	121.67
		N	122.33
		NE1	131.52
46	ALA	H	8.13
		HA	2.50
		HB1	-2.69
		HB2	-2.69
		HB3	-2.69
		CA	55.16
		CB	15.83
		N	119.28
47	HIS	H	12.62
		HA	8.18
		HB2	15.46
		HB3	18.03
		CA	63.44
		CB	24.83
47	HIS	N	122.07
48	LYS	H	9.45
		HA	5.17
		HB2	2.17
		HB3	2.65
		HG2	1.59
		HG3	1.99
		HD2	1.94
		HD3	2.06
		HE2	3.23
		CA	60.01
		CB	33.63
		CG	25.21
		CD	29.59
		CE	42.41
		N	120.47
49	THR	H	8.06
		HA	2.74
		HB	1.08
		QG2	-1.49
		CA	65.02
		CB	67.12
		CG2	19.50
		N	114.29
50	CYS	H	6.17
		HA	-2.99
		HB2	-0.80
		HB3	0.44
		CA	54.98
		CB	34.02
		N	117.73
51	LYS	H	6.18
		HA	5.73
		HB2	1.82
		HB3	3.43
		HD2	2.17
		HD3	2.91
		HE2	3.96
		CA	60.80
		CB	33.31
		CG	26.39

Table A2. 5 *(cont.)*

Residue number	Residue type	Atom	Chemical shift (ppm)
51	LYS	CD	30.65
		CE	42.29
		N	114.37
52	GLY	H	6.95
		HA2	4.18
		HA3	4.41
		CA	47.80
		N	103.81
53	CYS	H	6.78
		HA	4.27
		HB2	-0.20
		HB3	1.08
		CA	57.78
		CB	28.18
		N	119.73
54	HIS	H	10.38
		HA	9.04
		HB2	18.20
		HB3	21.62
		CA	60.57
		CB	13.50
		N	118.75
55	GLU	H	10.01
		HA	4.89
		HB2	2.64
		HB3	2.97
		HG2	2.95
		HG3	3.30
		CA	60.36
		CB	30.43
		CG	37.65
		N	119.90
56	GLU	H	8.58
		HA	4.30
		HB2	2.20
		HB3	2.31
		HG2	2.32
		HG3	2.46
		CA	59.68
		CB	30.21
		CG	35.97
56	GLU	N	121.31
57	MET	H	9.31
		HA	4.38
		HB2	2.25
		HB3	2.85
		HG2	2.53
		HE1	1.87
		HE2	1.87
		HE3	1.87
		CA	57.04
		CB	32.83
		CG	33.90
		CE	16.82
		N	114.65
58	LYS	H	8.30
		HA	4.33
		HB2	2.08
		HB3	2.22
		HG2	1.59
		HD2	1.80
		HD3	1.93
		HE2	3.18
		CA	57.37
		CB	29.17
		CG	25.18
		CD	29.24
		CE	42.64
		N	115.62
59	LYS	H	8.61
		HA	4.18
		HB2	0.93
		HB3	1.21
		HG2	1.02
		HG3	1.11
		HD2	1.20
		HE2	2.60
		CA	55.74
		CB	38.04
		CG	24.90
		CD	29.28
		CE	42.00

Table A2. 5 *(cont.)*

Residue number	Residue type	Atom	Chemical shift (ppm)
59	LYS	N	118.61
60	GLY	H	6.70
		HA2	-1.30
		CA	40.38
		N	103.47
61	PRO	HA	4.91
		HB2	1.47
		HB3	1.67
		HG2	-0.26
		HG3	1.70
		HD2	1.85
		CA	63.08
		CB	33.59
		CG	28.74
		CD	50.12
62	THR	H	9.41
		HA	4.96
		QG2	2.47
		CA	62.74
		CB	70.94
		CG2	23.65
		N	113.63
63	LYS	H	7.97
		HA	3.84
		HB2	1.29
		HB3	1.42
		HG2	1.48
		HD2	1.19
		HD3	1.36
		HE2	2.85
		CA	55.42
		CB	32.53
		CG	28.76
		CD	24.88
		CE	41.88
		N	123.77
64	CYS	H	6.68
		HA	-1.17
		HB2	-0.73
		HB3	1.15
		CA	55.56
64	CYS	CB	33.73
		N	117.16
65	GLY	H	8.45
		HA2	3.89
		HA3	5.26
		CA	45.80
		N	103.31
66	ASP	H	7.06
		HA	4.68
		HB2	2.19
		HB3	2.34
		CA	54.93
		CB	40.91
		N	116.80
67	CYS	H	6.84
		HA	5.43
		HB2	1.82
		HB3	3.03
		CA	57.33
		CB	39.48
		N	114.68
68	HIS	H	11.71
		HA	10.58
		HB2	9.15
		HB3	12.63
		CA	80.11
		CB	24.05
		N	121.43
69	LYS	H	10.54
		HA	5.96
		HB2	2.78
		HB3	2.96
		HG2	2.22
		HG3	2.28
		HD2	2.24
		HE2	3.31
		CA	57.01
		CB	35.15
		CG	25.60
		CD	29.63
		CE	42.84

Table A2. 5 *(cont.)*

Residue number	Residue type	Atom	Chemical shift (ppm)
69	LYS	N	127.70
70	LYS	H	9.20
		HA	5.08
		HB2	2.43
		HB3	2.49
		HG2	2.17
		HD2	2.23
		HE2	3.45
		CA	58.99
		CB	34.28
		CG	25.70
		CD	29.74
		CE	42.66
		N	130.18

Table A2. 6 – ^{1}H chemical shifts (ppm) of the heme substituents of PpcA V13 mutants in the reduced state, pH 8 and 15 ºC. The chemical shifts correspondent to the second conformation observed for V13I, V13S and V13T are given in parenthesis. PpcA values were previously determined [3] and are listed for comparison.

Heme substituent	Mutant	Heme I	Heme III	Heme IV
5H	V13A	9.64	10.59	9.03
	V13I	9.65 (9.62)	10.56 (10.51)	9.04 (9.04)
	V13S	9.65 (9.60)	10.54 (10.46)	9.05 (9.09)
	V13T	9.65 (9.57)	10.58 (10.45)	9.00 (9.11)
	PpcA	9.65	10.58	9.02
10H	V13A	9.10	9.85	9.32
	V13I	9.11 (9.02)	9.86 (9.86)	9.35 (9.26)
	V13S	9.00 (8.86)	9.85 (9.83)	9.43 (9.54)
	V13T	9.10 (8.84)	9.85 (9.85)	9.35 (9.54)
	PpcA	9.12	9.86	9.33
15H	V13A	9.27	9.47	9.54
	V13I	9.28 (9.28)	9.44 (9.44)	9.52 (9.52)
	V13S	9.31 (9.31)	9.43 (9.43)	9.52 (9.52)
	V13T	9.28 (9.28)	9.48 (9.48)	9.54 (9.54)
	PpcA	9.26	9.45	9.51
20H	V13A	9.52	10.12	9.40
	V13I	9.50 (9.50)	10.29 (10.34)	9.42 (9.42)
	V13S	9.49 (9.49)	10.32 (10.33)	9.46 (9.46)
	V13T	9.50 (9.50)	10.17 (10.38)	9.41 (9.41)
	PpcA	9.50	10.14	9.39
$2^{1}CH_3$	V13A	3.56	4.36	3.61
	V13I	3.56 (3.56)	4.41 (4.49)	3.64 (3.64)
	V13S	3.58 (3.58)	4.45 (4.60)	3.66 (3.66)
	V13T	3.58 (3.58)	4.39 (4.63)	3.64 (3.64)
	PpcA	3.56	4.35	3.61
$7^{1}CH_3$	V13A	3.57	4.14	3.03
	V13I	3.58 (3.56)	4.14 (4.14)	3.03 (3.03)
	V13S	3.56 (3.52)	4.14 (4.14)	3.04 (3.06)
	V13T	3.58 (3.50)	4.14 (4.14)	3.01 (3.07)
	PpcA	3.58	4.14	3.02
$12^{1}CH_3$	V13A	2.55	3.50	3.94
	V13I	2.55 (2.52)	3.51 (3.51)	3.95 (3.93)
	V13S	2.53 (2.50)	3.51 (3.51)	4.02 (4.04)
	V13T	2.57 (2.50)	3.51 (3.51)	3.99 (4.01)
	PpcA	2.55	3.50	3.95
$18^{1}CH_3$	V13A	3.35	3.86	3.36
	V13I	3.34 (3.34)	3.85 (3.84)	3.37 (3.37)
	V13S	3.33 (3.33)	3.84 (3.84)	3.41 (3.41)
	V13T	3.33 (3.33)	3.88 (3.79)	3.37 (3.37)

Table A2. 6 *(cont)*

Heme substituent	Mutant	Heme I	Heme III	Heme IV
$18^{1}CH_3$	PpcA	3.34	3.86	3.34
$3^{1}H$	V13A	6.28	6.94	6.03
	V13I	6.30 (6.27)	6.90 (6.89)	6.05 (6.05)
	V13S	6.28 (6.24)	6.91 (6.90)	6.05 (6.02)
	V13T	6.29 (6.19)	6.93 (6.89)	6.02 (6.06)
	PpcA	6.30	6.91	6.04
$8^{1}H$	V13A	6.27	6.59	6.30
	V13I	6.28 (6.19)	6.61 (6.61)	6.29 (6.25)
	V13S	6.18 (6.04)	6.61 (6.60)	6.33 (6.39)
	V13T	6.28 (6.00)	6.60 (6.60)	6.32 (6.43)
	PpcA	6.29	6.60	6.28
$3^{2}CH_3$	V13A	2.13	1.88	2.07
	V13I	2.14 (2.16)	1.73 (1.82)	2.06 (2.06)
	V13S	2.16 (2.19)	1.85 (2.00)	2.06 (2.04)
	V13T	2.14 (2.22)	1.82 (2.07)	2.06 (2.05)
	PpcA	2.14	1.73	2.06
$8^{2}CH_3$	V13A	1.81	2.97	1.55
	V13I	1.76 (1.70)	2.99 (2.99)	1.55 (1.55)
	V13S	1.69 (1.57)	2.98 (2.97)	1.58 (1.58)
	V13T	1.81 (1.55)	2.97 (2.97)	1.58 (1.60)
	PpcA	1.79	2.98	1.55

Table A2. 7 - Backbone chemical shifts (ppm) of reduced PpcA F6L, pH 7.1, 25 ºC. Atoms are indicated by the DYANA nomenclature [2]. Unassigned atoms are labeled N.A. The chemical shifts of the side chain of Leu[6] are also indicated.

Residue number	Residue type	Atom	Chemical shift (ppm)
1	ALA	H	N.A
		N	N.A
2	ASP	H	N.A
		N	N.A
3	GLU	H	7.63
		N	119.14
4	LEU	H	8.30
		N	123.01
5	THR	H	8.96
		N	120.17
6	LEU	H	9.65
		HA	5.04
		HB2	2.71
		HB3	2.43
		HG	2.61
		QD1	2.41
		QD2	2.28
		N	131.98
7	LYS	H	8.52
		N	124.97
8	ALA	H	8.55
		N	123.66
9	LYS	H	10.35
		N	125.41
10	ASN	H	9.01
		HD21	0.80
		HD22	5.57
		ND2	108.39
		N	113.89
11	GLY	H	7.02
		N	110.63
12	ASP	H	8.11
		N	129.52
13	VAL	H	8.70
		N	121.53
14	LYS	H	9.07
		N	129.65
15	PHE	H	9.26
		N	127.05
17	HIS	H	6.51
		N	132.65
17	HIS	HD1	8.72
		ND1	166.28
18	LYS	H	7.36
		N	119.38
19	LYS	H	7.32
		N	118.32
20	HIS	H	6.26
		HD1	8.03
		ND1	162.70
		N	114.07
21	GLN	H	7.37
		HE21	1.07
		NE2	105.20
		N	119.39
22	GLN	H	7.09
		HE21	6.70
		HE22	7.29
		NE2	111.35
		N	116.64
23	VAL	H	8.12
		N	117.86
24	VAL	H	8.72
		N	120.11
25	GLY	H	7.28
		N	105.90
26	ASN	H	7.76
		HD21	7.08
		HD22	7.91
		ND2	111.79
		N	116.97
27	CYS	H	9.01
		N	122.63
28	LYS	H	7.94
		N	115.15
29	LYS	H	7.53
		N	117.61
30	CYS	H	6.37
		N	111.71
31	HIS	H	7.17
		HD1	8.54
		ND1	163.96
		N	117.20

Table A2.7 *(cont)*

Residue number	Residue type	Atom	Chemical shift (ppm)
32	GLU	H	8.29
		N	125.46
33	LYS	H	8.29
		N	115.46
34	GLY	H	6.71
		N	107.39
36	GLY	H	3.84
		N	109.00
37	LYS	H	7.91
		N	116.94
38	ILE	H	9.94
		N	128.80
39	GLU	H	8.62
		N	131.66
40	GLY	H	9.16
		N	113.48
41	PHE	H	7.74
		N	120.40
42	GLY	H	6.10
		N	114.61
43	LYS	H	8.98
		N	122.68
44	ASP	H	7.89
		N	116.10
45	TRP	H	8.35
		HE1	10.27
		NE1	130.92
		N	121.49
46	ALA	H	8.61
		N	121.70
47	HIS	H	7.72
		N	117.09
48	LYS	H	6.98
		N	117.63
49	THR	H	8.32
		N	115.81
50	CYS	H	8.47
		N	121.75
51	LYS	H	5.98
		N	112.75
52	GLY	H	6.17
		N	101.25
53	CYS	H	7.07
		N	121.56
54	HIS	H	6.08
		HD1	9.64
		ND1	165.49
		N	117.33
55	GLU	H	7.93
		N	117.68
56	GLU	H	7.71
		N	120.79
57	MET	H	8.98
		N	115.14
58	LYS	H	7.75
		N	114.68
59	LYS	H	7.66
		N	118.88
60	GLY	H	7.62
		N	103.52
62	THR	H	8.74
		N	110.30
63	LYS	H	8.33
		N	124.48
64	CYS	H	8.76
		N	118.67
65	GLY	H	8.87
		N	102.28
66	ASP	H	7.73
		N	117.48
67	CYS	H	6.62
		N	115.66
68	HIS	H	6.96
		HD1	8.97
		ND1	166.25
		N	118.03
69	LYS	H	7.12
		N	124.15
70	LYS	H	7.91
		N	128.34

Table A2. 8 - Backbone chemical shifts (ppm) of reduced PpcA W45M, pH 7.1, 25 ºC. Atoms are indicated by the DYANA nomenclature [2]. Unassigned atoms are labeled N.A. The chemical shifts of the tentative assignment of the side chain of Met[45] are also indicated.

Residue number	Residue type	Atom	Chemical shift (ppm)
1	ALA	H	N.A
		N	N.A
2	ASP	H	N.A
		N	N.A
3	GLU	H	7.55
		N	118.94
4	LEU	H	8.08
		N	123.71
5	THR	H	8.43
		N	117.31
6	PHE	H	9.56
		N	128.99
7	LYS	H	9.07
		N	125.71
8	ALA	H	8.42
		N	124.84
9	LYS	H	10.37
		N	125.65
10	ASN	H	9.08
		HD21	0.88
		HD22	5.63
		ND2	108.28
		N	113.70
11	GLY	H	7.05
		N	110.52
12	ASP	H	8.05
		N	129.30
13	VAL	H	8.68
		N	120.85
14	LYS	H	8.96
		N	128.91
15	PHE	H	9.35
		N	125.92
17	HIS	H	6.52
		HD1	8.70
		ND1	166.21
		N	132.82
18	LYS	H	7.37
		N	119.15
19	LYS	H	7.22
		N	118.38
20	HIS	H	6.26
		HD1	7.98
		ND1	162.78
		N	114.05
21	GLN	H	7.37
		HE21	1.04
		NE2	105.12
		N	119.39
22	GLN	H	7.10
		HE21	6.71
		HE22	7.29
		NE2	111.32
		N	116.62
23	VAL	H	8.04
		N	117.94
24	VAL	H	8.75
		N	120.22
25	GLY	H	7.32
		N	105.78
26	ASN	H	7.71
		HD21	7.06
		HD22	7.89
		ND2	111.95
		N	117.01
27	CYS	H	8.96
		N	122.45
28	LYS	H	7.92
		N	115.18
29	LYS	H	7.52
		N	117.72
30	CYS	H	6.31
		N	111.59
31	HIS	H	7.09
		HD1	8.60
		ND1	163.97
		N	116.86
32	GLU	H	8.17
		N	125.68

Table A2. 8 *(cont)*

Residue number	Residue type	Atom	Chemical shift (ppm)
33	LYS	H	8.23
		N	115.19
34	GLY	H	6.67
		N	107.20
36	GLY	H	3.84
		N	109.34
37	LYS	H	8.08
		N	116.65
38	ILE	H	10.62
		N	128.85
39	GLU	H	8.69
		N	130.88
40	GLY	H	9.22
		N	113.49
41	PHE	H	8.04
		N	121.70
42	GLY	H	5.98
		N	114.54
43	LYS	H	8.92
		N	122.31
44	ASP	H	7.87
		N	115.40
45	MET	H	8.03
		HA	3.92
		HB2	3.23
		HB3	2.69
		HG2	2.50
		HG3	2.10
		N	119.18
46	ALA	H	8.77
		N	121.70
47	HIS	H	7.24
		N	115.98
48	LYS	H	7.08
		N	117.98
49	THR	H	8.68
		N	114.98
50	CYS	H	8.39
		N	123.06
51	LYS	H	5.97
		N	112.78
52	GLY	H	6.41
		N	101.72
53	CYS	H	7.09
		N	121.48
54	HIS	H	6.09
		HD1	9.62
		ND1	165.56
		N	117.41
55	GLU	H	8.01
		N	117.68
56	GLU	H	7.70
		N	120.82
57	MET	H	9.00
		N	115.15
58	LYS	H	7.78
		N	114.80
59	LYS	H	7.68
		N	118.85
60	GLY	H	7.62
		N	103.41
62	THR	H	8.76
		N	110.62
63	LYS	H	8.38
		N	124.55
64	CYS	H	8.76
		N	118.74
65	GLY	H	8.87
		N	102.37
66	ASP	H	7.75
		N	117.65
67	CYS	H	6.64
		N	115.71
68	HIS	H	7.02
		HD1	8.89
		ND1	166.50
		N	118.06
69	LYS	H	7.28
		N	124.24
70	LYS	H	7.93
		N	128.60

Table A2. 9 - ^{1}H chemical shifts (ppm) of the heme substituents of PpcA F6L and PpcA W45M in the reduced state, pH 7.1, 25 ºC. Heme substituents are indicated in the IUPAC-IUB nomenclature for tetrapyrroles [1].

Heme Substituent	Mutant	Heme I	Heme III	Heme IV
5H	F6L	9.42	9.76	9.01
	W45M	9.40	9.74	9.03
10H	F6L	9.16	9.84	9.35
	W45M	9.13	9.86	9.35
15H	F6L	9.22	9.51	9.50
	W45M	9.32	9.50	9.44
20H	F6L	9.51	10.15	9.43
	W45M	9.51	10.22	9.43
$2^{1}CH_3$	F6L	3.53	4.34	3.62
	W45M	3.53	4.37	3.61
$7^{1}CH_3$	F6L	3.50	3.94	3.01
	W45M	3.45	4.02	3.04
$12^{1}CH_3$	F6L	2.67	3.46	3.90
	W45M	2.67	3.48	3.67
$18^{1}CH_3$	F6L	3.32	3.91	3.33
	W45M	3.35	3.93	3.34
$3^{1}H$	F6L	5.97	6.71	5.95
	W45M	5.96	6.54	5.92
$8^{1}H$	F6L	6.40	6.57	6.31
	W45M	6.27	6.57	6.34
$3^{2}CH_3$	F6L	2.11	2.46	2.06
	W45M	2.11	2.16	2.06
$8^{2}CH_3$	F6L	1.99	2.93	1.5
	W45M	1.87	2.95	1.54

A.3 Deduction

Deduction of the equation used in the fitting of the experimental data from redox titrations followed by UV-Visible spectroscopy [4].

The reduction potential of a monoheme cytochrome (E^0) can be obtained by the direct application of the Nernst Equation (Equation A1) because only the fully reduced and oxidized states coexist in solution.

$$E = E^0 + \frac{RT}{nF} ln \frac{P_{ox}}{P_{red}} \quad \text{(A1)}$$

In Equation A1, E is solution potential, n the number of electrons involved in the reaction, R the gas constant, T the temperature, F the Faraday constant and P_{ox} and P_{red} the molar fractions of the reduced and oxidized species in solution.

In the case of a triheme cytochrome, four macroscopic oxidation stages – S_0 (fully reduced), S_1 (containing all the microstates with one oxidized heme), S_2 (containing all the microstates with two oxidized hemes) and S_3 (fully oxidized) – are related to each other by three reversible one-electron transfer steps that convert the fully reduced in the fully oxidized state (see Figure 1.6):

$$S_0(P_0) \underset{E_1}{\rightleftharpoons} S_1(P_1) \underset{E_2}{\rightleftharpoons} S_2(P_2) \underset{E_3}{\rightleftharpoons} S_3(P_3)$$

In the previous scheme P_0, P_1, P_2 and P_3 correspond to the molar fractions of the microstates in oxidation stages S_0, S_1, S_2 and S_3, respectively. Hence, the macroscopic redox behavior of the protein can be characterized by combining three Nernst equations that relate the four one-electron macroscopic transfer steps (Equations A2-A4):

$$E = E_1 + \frac{RT}{F} ln \frac{P_1}{P_0} \Leftrightarrow P_1 = P_0 e^{(E-E_1)\frac{F}{RT}} \quad \text{(A2)}$$

$$E = E_2 + \frac{RT}{F} ln \frac{P_2}{P_1} \Leftrightarrow P_2 = P_0 e^{(2E-E_1-E_2)\frac{F}{RT}} \quad \text{(A3)}$$

$$E = E_3 + \frac{RT}{F} ln \frac{P_3}{P_2} \Leftrightarrow P_3 = P_0 e^{(3E-E_1-E_2-E_3)\frac{F}{RT}} \quad \text{(A4)}$$

In Equations A2-A4, E, E_1, E_2 and E_3 represent, respectively, the solution potential measured experimentally and the macroscopic reduction potential for each oxidation step.

The reduced fraction is described by the ratio between the number of reduced species and the total number of species, normalized for the three heme groups, as described by Equation A5.

$$\text{Reduced fraction} = \frac{3P_0 + 2P_1 + P_2 + 0P_3}{3P_0 + 2P_1 + P_2 + 0P_3 + 0P_0 + P_1 + 2P_2 + 3P_3} \Leftrightarrow \tag{A5}$$

$$\text{Reduced fraction} = \frac{3P_0 + 2P_1 + P_2 + 0P_3}{3(P_0 + P_1 + P_2 + P_3)}$$

Substituting the P_1, P_2 and P_3 variables in Equation A5 by the corresponding Nernst expressions (Equations A2-A4), we obtain the final expression used for the experimental data fitting of the UV-Visible redox titrations performed in this work indicated in Equation A6:

$$\text{Reduced fraction} = \frac{3 + 2e^{\left[(E-E_1)\frac{F}{RT}\right]} + e^{\left[(2E-E_1-E_2)\frac{F}{RT}\right]}}{3(1 + e^{\left[(E-E_1)\frac{F}{RT}\right]} + e^{\left[(2E-E_1-E_2)\frac{F}{RT}\right]} + e^{\left[(3E-E_1-E_2-E_3)\frac{F}{RT}\right]}} \tag{A6}$$

A.4 References

[1] G.P. Moss, Nomenclature of tetrapyrroles, European Journal of Biochemistry, 178 (1988) 277-328.

[2] P. Güntert, C. Mumenthaler, K. Wüthrich, Torsion angle dynamics for NMR structure calculation with the new program DYANA, Journal of Molecular Biology, 273 (1997) 283-298.

[3] L. Morgado, M. Bruix, V. Orshonsky, Y.Y. Londer, N.E. Duke, X. Yang, P.R. Pokkuluri, M. Schiffer, C.A. Salgueiro, Structural insights into the modulation of the redox properties of two *Geobacter sulfurreducens* homologous triheme cytochromes, Biochim. Biophys. Acta, 1777 (2008) 1157-1165.

[4] D.L. Turner, C.A. Salgueiro, T. Catarino, J. Legall, A.V. Xavier, NMR studies of cooperativity in the tetrahaem cytochrome c_3 from *Desulfovibrio vulgaris*, European Journal of Biochemistry, 241 (1996) 723-731.